

Universitat Politècnica de Catalunya  
Facultat de Matemàtiques i Estadística

Master in Advanced Mathematics and Mathematical Engineering  
MASTER'S THESIS

**Agent-based modelling of  
embryonic organoid development**

**Víctor Villegas-Morral**

Supervised by David Oriola Santandreu  
September 2024



UNIVERSITAT POLITÈCNICA DE CATALUNYA  
BARCELONATECH

Facultat de Matemàtiques i Estadística



## Acknowledgements

I want to thank David Oriola Santandreu for his guidance throughout this project, as well as the entire BIOCOP-SC team for welcoming me over these past few months. I also want to express my gratitude to Gabriel Torregrosa Cortés for his help and advice on the development of the code. Finally, I would want to toast all the wonderful people I have met during the master's. I wish to continue seeing all of you – even if it's not on a Thursday. Wishing everyone the best. Moltes gràcies.

## Abstract

In recent years, embryonic organoids have become an increasingly important tool for studying complex developmental processes that are difficult to access in actual animal embryos. These *in vitro* structures can be generated in a high-throughput manner and have the advantage of mimicking many intricate behaviours observed in real embryos. In this study, we describe an agent-based model designed to simulate the development of embryonic organoids, and implement it using CellBasedModels.jl, the Julia package for multicellular modelling. Our model integrates mechanical interactions with cellular differentiation and accounts for the stochastic nature of both cell-cell interactions and cell fate transitions. We provide a comprehensive overview of the key simplifications made in the mathematical model and offer a numerical justification for these choices. Finally, we use the model to simulate the early stages of gastruloid formation, analysing the roles of cell differentiation and differential adhesion.

## Resum

En els darrers anys, els organoides embrionaris han esdevingut una eina cada cop més important per estudiar processos de la biologia del desenvolupament difícils d'explorar en embrions animals. Aquestes estructures *in vitro* són fàcilment reproduïbles i tenen l'avantatge de mimetitzar molts dels comportaments complexos observats en embrions reals. En aquest estudi, contruïm un model basat en agents per simular el desenvolupament d'organoides embrionaris i l'implementem utilitzant CellBasedModels.jl, el paquet de Julia de modelització multicel·lular. El nostre model integra les interaccions mecàniques amb la diferenciació cel·lular, i té en compte la naturalesa estocàstica tant de les interaccions entre cèl·lules com de la diferenciació cel·lular. Presentem una visió completa de les principals simplificacions realitzades en el model matemàtic i oferim una justificació numèrica per aquestes eleccions. Finalment, utilitzem el model per simular les primeres etapes de la formació de *gastruloids*, analitzant els rols de la diferenciació cel·lular i l'adhesió diferencial.

# Contents

<b>Nomenclature</b>	<b>1</b>
<b>1 Introduction</b>	<b>3</b>
1.1 Embryonic development and organoids . . . . .	3
1.2 Modelling of organoids . . . . .	4
1.3 Structure of the thesis . . . . .	6
<b>2 Agent-based modelling of 3D cellular aggregates</b>	<b>7</b>
2.1 Proliferation . . . . .	7
2.2 Mechanical interactions . . . . .	8
2.2.1 The equations of motion . . . . .	8
2.2.2 Accounting for random cell-cell interactions . . . . .	10
2.2.3 Simplification of the mechanics . . . . .	12
2.3 Cell differentiation . . . . .	13
2.3.1 Linear first order differentiation kinetics . . . . .	15
2.3.2 Nonlinear differentiation kinetics . . . . .	15
2.4 Nondimensionalization . . . . .	16
<b>3 Implementation of the model</b>	<b>18</b>
3.1 Computational framework and numerical methods . . . . .	18
3.1.1 Framework . . . . .	18
3.1.2 Integration of the equations of motion . . . . .	19
3.1.3 Algorithm for cell differentiation . . . . .	20
3.2 Numerical validation of the global friction approximation . . . . .	20
3.2.1 Aggregate stability . . . . .	20
3.2.2 Global friction and average number of neighbours . . . . .	21
3.2.3 Absence of net flows . . . . .	26
<b>4 Test examples and applications</b>	<b>29</b>

4.1	Physical description . . . . .	30
4.1.1	Experimental system . . . . .	30
4.1.2	Proliferation speed . . . . .	30
4.1.3	Protrusion force . . . . .	31
4.1.4	Measure of movement . . . . .	32
4.1.5	Stable timesteps . . . . .	34
4.2	Effect of differentiation kinetics on the evolution of the fate proportions . . . . .	34
4.2.1	Differentiation without proliferation . . . . .	35
4.2.2	Differentiation with proliferation . . . . .	37
4.3	Proportion of B cells in terms of b . . . . .	38
4.3.1	Using cell-cell feedback . . . . .	40
4.4	Differential cell adhesion . . . . .	40
4.4.1	Differential proliferation . . . . .	42
5	Conclusions and further work	45
	<b>Appendix</b>	<b>48</b>
A.1	Mathematical derivations . . . . .	48
A.2	Code . . . . .	51
A.2.1	Technical specifications . . . . .	51
A.2.2	Programs . . . . .	51
A.2.3	Packages . . . . .	52
A.3	Additional figures . . . . .	53
	<b>References</b>	<b>55</b>

# Nomenclature

This section serves as a reference for the notation used throughout the text. Quantities are presented only in their dimensionless form, denoted with a  $\sim$ . Physical values can be recovered using the methods presented in Section 2.4.

Symbol	Code	Description	Value
$T_0$	t0	Typical time scale	2 h
$R_0$	r0	Typical length scale	5 $\mu\text{m}$
$F_0$	f0	Typical force scale	20 pN
$\tilde{t}$	t	Simulation time	
$\Delta\tilde{t}$	dt	Timestep	0.002
$N$	N	Total number of cells	
$\tilde{r}$	r	Radius	1
$\alpha_{\text{ov}}$	olap	Overlap factor	0.75
$\tilde{t}_{\text{div}}$	t_div	Division time	
$\tilde{\tau}_{\text{div}}$	tau_div	Average division time	5–10
	x_div, y_div, z_div	Division axis vector	
$\tilde{\sigma}_{\text{div}}$	sigma_div	Division time variability	0.5
$\text{U}(a, b)$		Uniform distribution	
	d_on	Differentiation trigger	
	g_on	Growth trigger	
$m$		Mass	
$\tilde{x} \in \mathbb{R}^3$	x, y, z	Position	
$\tilde{v} \in \mathbb{R}^3$	vx, vy, vz	Velocity	
$\tilde{\lambda}$	lambda	Global friction coefficient	1
$\tilde{\Lambda}$	lambda	Variable fiction coefficient	0.1

$\tilde{F}_i \in \mathbb{R}^3$	<code>fx, fy, fz</code>	Sum of passive forces	
$\tilde{F}_{ij} \in \mathbb{R}^3$		Force exerted by cell $j$ over cell $i$	
$\tilde{d}_{ij}$	<code>d_ij</code>	Euclidean distance between cells $i$ and $j$	
$\alpha_{\text{nb}}$	<code>range</code>	Neighbour range factor	1.2
$\mu$	<code>mu</code>	Force range factor	2
$n_i$	<code>ni</code>	Number of neighbours	
	<code>ni_a</code>	Number of neighbours in state $A$	
$U_i$		Set of neighbours	
$\alpha_{\text{adh}}$	<code>f_adh</code>	Adhesion factor	1
$\alpha_{\text{rep}}$	<code>rep</code>	Repulsion factor	2.5
$\tilde{F}_p \in \mathbb{R}^3$	<code>fpx, fpy, fpz</code>	Protrusion force vector	
$\tilde{D}$	<code>fp</code>	Protrusion strength	10
$P_i$		Protrusion factor	
	<code>marked</code>	Protrusion marker	
$\tilde{t}_{ij}$	<code>t_paired</code>	Protrusion time	10
$\tilde{k}_{p_{\text{on}}}$	<code>k_p_on</code>	Protrusion activation rate	20
$\tilde{k}_{p_{\text{off}}}$	<code>k_p_off</code>	Protrusion deactivation rate	10
$A, B, C$	1, 2, 3	Cell states	
$\text{state}(i)$	<code>cell_state</code>	State of the cell	
$\mathbb{1}_X(i)$		State indicator function	
$\phi_X$	<code>phi_x</code>	Proportion of cells in state $X$	
$\phi_B(0)$	<code>b</code>	Initial proportion of cells in state $B$	0.2
$\langle \mathbb{1}_X(i) \rangle$		Proportion of neighbours in state $X$	
$\tilde{r}_{X \rightarrow Y}$	<code>r_ab, r_bc</code>	Transition rate from $X$ to $Y$	
$\tilde{p}$	<code>p</code>	Fixed transition rate from $A$ to $B$	0.25
$\tilde{q}$	<code>q</code>	Fixed transition rate from $B$ to $C$	$0.5\tilde{p}$
$K$	<code>k</code>	Inhibition factor	4.76
$d_2(t_n)$	<code>msd</code>	Mean square displacement	
$g_2(t_n)$		Mean square displacement step	

# Chapter 1

## Introduction

Biological systems are self-organized at various scales, from forests down to molecules. In particular, the origin of animal lifeforms occurs through embryonic development, where stem cells divide and differentiate to form the different tissues that will constitute the organism.

### 1.1 Embryonic development and organoids

During development, pluripotent stem cells (PSCs) initiate the process of cellular differentiation, eventually forming all the tissues in the adult animal. Advances in stem cell biology have enabled the generation of *in vitro* models of morphogenesis from PSCs, forming organ-like structures called organoids (Huch et al. 2017) (see Figure 1.1). The emerging field of synthetic embryology has used the previous tools to investigate in a dish previously inaccessible processes that occur in embryos (Liu and Fletcher 2009; Serra et al. 2019; Gritti, Oriola, and Trivedi 2021; Sullivan and Santos 2023).

In particular, recent studies have succeeded into recapitulating gastrulation, the early stage of embryonic development in which the embryo transforms into a multilayered structure. Such organoid models are known as gastruloids (see Figure 1.2) (Brink et al. 2014; Simunovic and Brivanlou 2017). These organoids form structures similar to those unique to mammalian embryos that define the body axis, and further research is expected to reveal yet unknown mechanisms of development (Turner 2017).

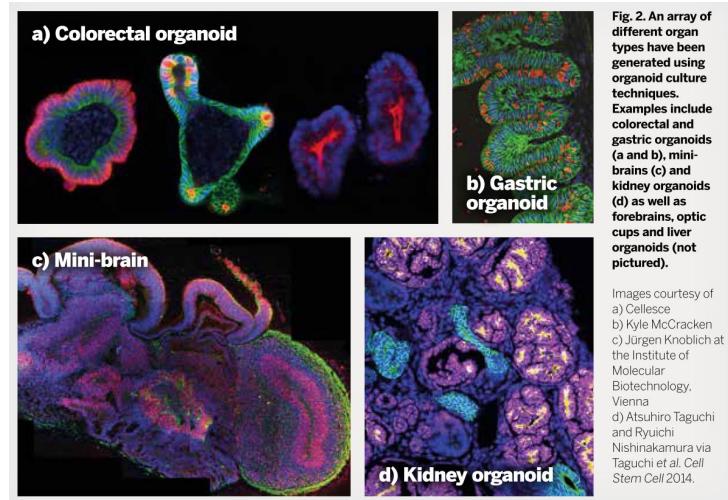


Figure 1.1: Organoids reproducing different organs. Adapted from Turner 2017.

## 1.2 Modelling of organoids

During embryogenesis, cells communicate and self-organize to form multicellular structures, and cell behaviour arises from both physical and chemical interactions (see Fig 1.3). On the one hand, adhesion, repulsion, and attractive forces influence cell movement. On the other hand, cell-cell signalling drives cell differentiation, which in turn affects cell motion. Experimental setups for organoid generation typically involve culturing cells in small volumes under controlled conditions that allow for the monitoring of variables such as cell proliferation, differentiation, and movement.

One of the key aspects in embryonic development is how the embryo breaks symmetry to form the different body axes. Recent work using the gastruloid system suggests that the timing of this process is controlled by cell-cell com-

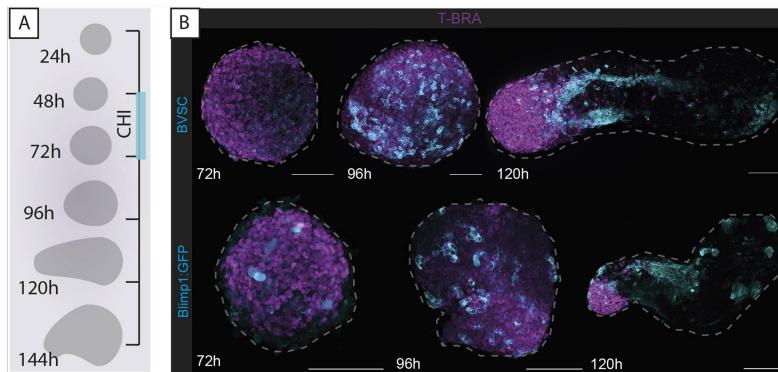


Figure 1.2: Morphological changes of a gastruloid over time, using markers for different cell states. Adapted from Cooke et al. 2023.

munication, and depends on the initial proportions of different type of cells (Oriola, Torregrosa-Cortés, et al. [unpublished](#)). We aim to construct a model that simulates the early stages of 3D gastruloid development in order to analyse cell proportions and differentiation kinetics.

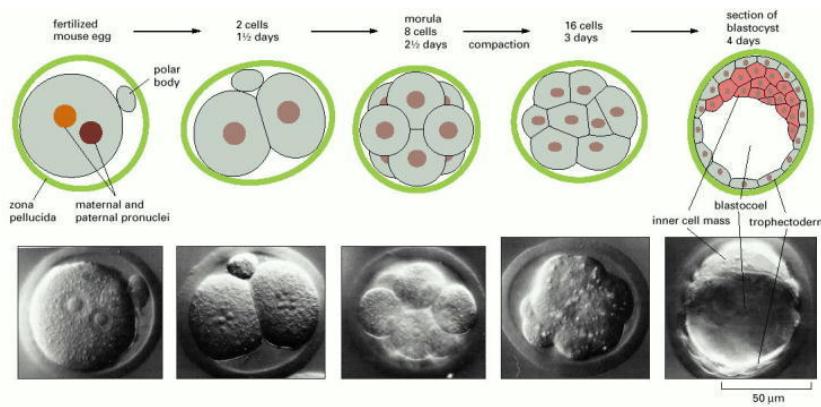


Figure 1.3: Early stages of mouse embryonic development. Adapted from Alberts, Johnson, and Lewis 2002.

Computational models make it possible to study the interplay between cell mechanics and cell signalling. Depending on the focus of analysis, organoid models are divided into two main categories: continuum models and agent-based models (ABMs). The former describe tissues as a whole using partial differential equations, while ABMs represent tissue dynamics as the interplay of individual agents (Van Liedekerke et al. 2015).

In this study, we will focus on centre-based models, representing agents as spheres. Cells can divide, move, and interact with other cells, with their motion described by a set of equations for each cell based on pairwise forces. This model supports precise tracking of agents, heterogeneity in cell types, a well-defined timescale, and the modelling of cell-cell interactions (Van Liedekerke et al. 2015; Gritti, Oriola, and Trivedi 2021).

Computations are carried out using the Julia (Bezanson et al. 2017) package CellBasedModels.jl (Torregrosa-Cortés et al. [unpublished](#)). Julia is an open source dynamic programming language for scientific computing, characterized by its high performance and attainable reproducibility. The CellBasedModels.jl package offers an understandable and efficient framework for the fast simulation of multicellular communities, allowing mechanical and biochemical interactions to be modelled.

## 1.3 Structure of the thesis

Expanding on the work of Van Liedekerke et al. 2015, Saiz et al. 2020, and Torregrosa-Cortés 2023, we develop and implement a model that aims to simulate gastruloids by coupling mechanical interactions and cell fate decisions within a stochastic framework. In Chapter 2, we present the mathematical model, describing cell mechanics and cell differentiation. In Chapter 3, we review the methods used to implement the model, discussing key simplifications necessary for numerical stability. Finally, Chapter 4 presents the applications and analyses performed using the simulation motivated by the results presented in Oriola, Torregrosa-Cortés, et al. *unpublished*. The Appendix includes mathematical derivations, brief descriptions of the programs used, and supplementary figures.

The primary aim of this work is to design an agent-based model that simulates stable 3D multicellular aggregates, providing clear explanations behind each choice in the model, and to contribute to the scientific community by expanding the documentation for the CellBasedModels.jl package.

All the code developed for this project is available on the [Multiscale Physics of Living Systems Group's GitHub](#)<sup>1</sup> and on my [personal GitHub](#)<sup>2</sup>.

---

<sup>1</sup><https://github.com/MPoLS-lab/villegas-morral-msc-thesis>

<sup>2</sup><https://github.com/villegas-morral/masters-thesis>

# Chapter 2

## Agent-based modelling of 3D cellular aggregates

This chapter describes the behaviour of the cellular aggregates studied in this work, and presents some simplifications to allow for its simulation. The aggregates are modelled as systems of identical spherical cells that transition from one state to another using an agent-based model.

The movement of each cell in the aggregate is governed by its equations of motion. Those depend on the interactions of each cell with nearby cells and account for friction, adhesion, and active forces. Cell differentiation is described by transition rates, which may depend on the state of the system. Our aim is to couple mechanics and cell fate transitions keeping the stochastic behaviour of both cell motion and differentiation.

### 2.1 Proliferation

We start from a single cell that divides into two daughter cells of the same volume, thus increasing the total volume of the aggregate over time (see Figure 2.1).

**Definition 1.** Consider an initial cell of radius  $r$ , null velocity, and an arbitrary division time  $t_{div}$ . The rules for division for an arbitrary cell  $i$  are:

1. When the simulation time  $t$  exceeds the division time  $t_{div_i}$ , the cell divides into two identical cells positioned in opposite directions. The division axis is chosen uniformly at random and, the separation between the centres of the cells is  $2r\alpha_{ov}$ ,

where  $0 < \alpha_{ov} \leq 1$ .

2. A new division time is assigned at random to each daughter cell using the following expression,

$$t_{div_i} \leftarrow t + U(\tau_{div}(1 - \sigma_{div}), \tau_{div}(1 + \sigma_{div})),$$

where  $U(a, b)$  denotes the uniform distribution. Both daughter cells are initialized at half of the velocity of the mother cell before division in order to enforce momentum conservation.

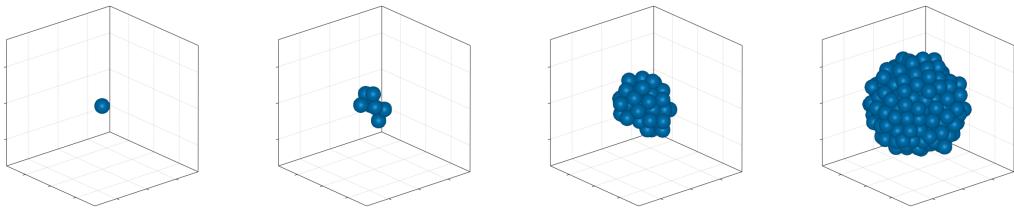


Figure 2.1: Formation of a 300-cell multicellular aggregate.

## 2.2 Mechanical interactions

### 2.2.1 The equations of motion

Cells exert and receive forces that condition their movement and organization in the aggregate. Next, we define the equations of motion of each cell, considering friction forces, passive forces, and active forces (Van Liedekerke et al. 2015). Let us define the system and break down every component.

**Definition 2.** Each cell  $i$  behaves according to the following equations of motion,

$$\begin{aligned} m \frac{dv_i}{dt} &= -\Lambda \sum_{j \in U_i} (v_i - v_j) + \sum_{j=1}^N F_{ij} + F_{a_i} \\ \frac{dx_i}{dt} &= v_i, \end{aligned} \tag{2.1}$$

with initial conditions

$$v_i(0) = v_{0_i} \quad x_i(0) = x_{0_i}, \tag{2.2}$$

where  $x, v \in \mathbb{R}^3$  represent the position of the centre and the velocity, respectively. Here,  $\Lambda$  is the friction coefficient, and  $U_i$  denotes the set of neighbouring cells.

**Definition 3.** Two cells  $i$  and  $j$  are said to be neighbours if and only if

$$d_{ij} < 2\alpha_{nb}r,$$

where  $\alpha_{nb} \in \mathbb{R}$  is the neighbouring range, and  $d_{ij} = |x_i - x_j|$  is the Euclidean distance in  $\mathbb{R}^3$  between cells  $i$  and  $j$ . The set of neighbours (or neighbourhood) of cell  $i$  is defined as

$$U_i = \{j : d_{ij} < 2\alpha_{nb}\},$$

and  $n_i = |U_i|$  denotes the number of neighbours of cell  $i$ .

**Notation 1.** Variables regarding the agents change over time. Notation such as  $U_i$  and  $d_{ij}$  refers to the time-dependent functions  $U_i(t)$  and  $d_{ij}(t)$ .

The first term in the right-hand side of Equation 2.1 represents the friction forces. These forces are dependent on the relative velocity of the cell with respect to its neighbours. The last two terms correspond to passive and active forces, respectively. Active forces can be due to processes such as filopodia between cells, and will be described later in this section.

Passive forces exerted by nearby cells are modelled as piecewise functions with repulsive and attractive regimes, which take into account steric interactions and cell-cell adhesion, respectively (Saiz et al. 2020; Torregrosa-Cortés 2023).

**Notation 2.** The range of cells that forces take into account is slightly larger than the neighbouring range. These are referred to as nearby cells throughout the text to distinguish them from neighbouring cells.

**Definition 4.** The passive force  $F_{ij} \in \mathbb{R}^3$  exerted by cell  $j$  over cell  $i$  is given by

$$F_{ij} = \begin{cases} \frac{(x_i - x_j)}{d_{ij}} F_0 f_r(d_{ij}) \alpha_{adh} \alpha_{rep} & \text{if } d_{ij} \leq 2r \\ \frac{(x_i - x_j)}{d_{ij}} F_0 f_r(d_{ij}) \alpha_{adh} & \text{if } 2r < d_{ij} < 2\mu r \\ 0 & \text{otherwise,} \end{cases} \quad (2.3)$$

where  $\mu > 0$  is the force range factor, and  $\alpha_{adh}, \alpha_{rep} > 0$  are the adhesion and repulsion factors, respectively. Here,  $F_0 > 0$  denotes the typical force generated by cells, and

$$f_r(d_{ij}) = \left( \frac{2r}{d_{ij}} - 1 \right) \left( \frac{2\mu r}{d_{ij}} - 1 \right) \in \mathbb{R}. \quad (2.4)$$

The reasoning behind this expression is discussed in Section 3.2.2.

**Remark 1.** *The adhesion factor accounts for differences in the adhesion strength for each pair. Unless stated otherwise, it is assumed to be  $\alpha_{adh} \equiv 1$ .*

The function  $f_r$  determines whether the force is attractive ( $f_r < 0$ ) or repulsive ( $f_r > 0$ ). When there exists an overlap between cells, e.g. after the two daughter cells are created, cells are pushed away from each other until their distance reaches the equilibrium distance  $2r$ .

**Remark 2.** *The behaviour of the force depends on the distance as follows.*

- *If  $d_{ij} < 2r$ , then  $f_r > 0$ . The force is repulsive when the cells are separated below their equilibrium distance.*
- *If  $d_{ij} = 2r$ , then  $f_r = 0$ . There is no force when the separation is exactly the equilibrium distance.*
- *If  $2r < d_{ij} < 2\mu r$ , then  $f_r < 0$ . The force is attractive when the cells are separated slightly above the equilibrium distance.*
- *If  $2\mu r \leq d_{ij}$ , then  $f_r = 0$ . There is no force when the cells are far apart.*

A sketch of the  $f_r$  function is shown in Figure 2.2.

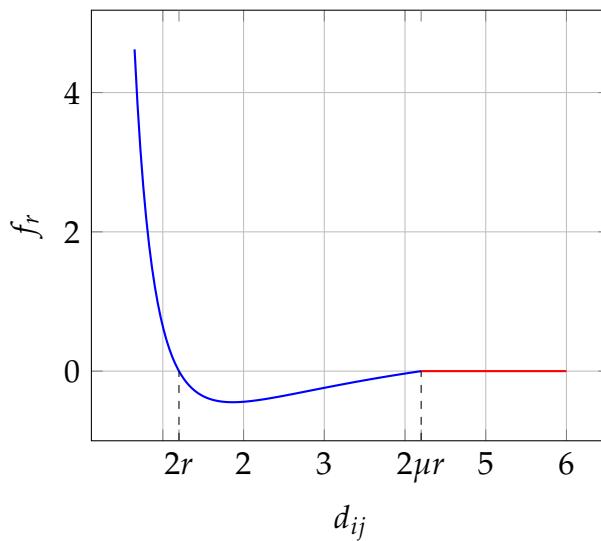


Figure 2.2: Example plot of the function  $f_r$  for  $r = 0.6$ ,  $\mu = 3.5$ .

### 2.2.2 Accounting for random cell-cell interactions

Cells exhibit subtle displacements caused by active forces, causing a continuous reorganization of the aggregate. We model such forces to mimic the behaviour

of filopodia, dynamic structures that extend from the cell membrane and connect to nearby cells. These cell-cell interactions result in temporal directed pairwise forces, which we refer to as protrusions (Oriola, Marin-Riera, et al. 2022; Torregrosa-Cortés 2023).

**Definition 5.** *Each cell has an associated protrusion function  $P_i \in \{0, 1, -1\}$  such that  $P_i = 0$  if the protrusion is inactive. Protrusions are modelled according to the following rules:*

1. *The probability of activating an inactive protrusion after a time  $\Delta t$  is  $k_{p_{on}}\Delta t$ , where  $k_{p_{on}}$  is the protrusion activation rate.*
2. *If  $P_i$  is activated at the previous step, set  $P_i = 1$ .*
  - i. *Choose a random inactive nearby cell  $j$  and set  $P_j = -1$ .*
  - ii. *Choose a random number  $\xi \sim U(0, 1)$ .*
  - iii. *Let  $k_{p_{off}}$  be the protrusion deactivation rate. Set a random protrusion time  $t_{ij}$  to both  $i$  and  $j$ ,*
$$t_{ij} \leftarrow t - \frac{\log(\xi)}{k_{p_{off}}}.$$
3. *While  $t < t_{ij}$ , cell  $i$  exerts a protrusion force  $F_{p_i} \in \mathbb{R}^3$  on cell  $j$ .*
4. *When  $t \geq t_{ij}$ , then  $P_i = P_j = 0$  and the cells become again eligible to be connected with another random inactive cell.*

**Definition 6.** *Let  $D > 0$  be the protrusion strength constant. The protrusion force is given by*

$$F_{p_i} = \begin{cases} -\frac{(x_i - x_j)}{d_{ij}} D \alpha_{rep} & \text{if } d_{ij} \leq 2r \\ -\frac{(x_i - x_j)}{d_{ij}} D & \text{if } 2r < d_{ij} < 2\mu r, \end{cases} \quad (2.5)$$

and  $|F_{p_i}| \in \{D\alpha_{rep}, D\}$ .

**Remark 3.** *Modelling the active forces as the protrusion force, Equation 2.1 becomes*

$$\begin{aligned} m \frac{dv_i}{dt} &= -\Lambda \sum_{j \in U_i} (v_i - v_j) + \sum_{j=1}^N F_{ij} + P_i F_{p_i} \\ \frac{dx_i}{dt} &= v_i, \end{aligned} \quad (2.6)$$

where  $P_i F_{p_i} \in \{-F_{p_i}, F_{p_i}, 0\}$ .

### 2.2.3 Simplification of the mechanics

Next, we simplify the equations of motion that govern the model. In the context of cell movement in cellular tissues, viscous forces dominate over inertial forces; thus, the system is assumed to be in the overdamped regime (Purcell 1977; Van Liedekerke et al. 2015).

Under these circumstances, the inertia term  $m \frac{dv_i}{dt}$  is negligible, and thus the previous equation becomes

$$0 = -\Lambda \sum_{j \in U_i} (v_i - v_j) + \sum_{j=1}^N F_{ij} + P_i F_{p_i}. \quad (2.7)$$

In 2D cell monolayers, the friction force originates from the interaction between the cell and the substrate, and is proportional to the velocity of the cell. However, in 3D aggregates the friction force is proportional to the relative velocities between cells (Van Liedekerke et al. 2015), as stated in Equation 2.7. Still, we will show that, performing the appropriate approximations, one can simplify the friction term to  $\Lambda v_i$ .

In Ya | l a (Germann, Marin-Riera, and Sharpe 2019), a CUDA/C++ agent-based model, the friction term in Equation 2.7 is simplified as follows,

$$-n_i \Lambda \left( v_i - \frac{1}{n_i} \sum_{j \in U_i} v_j \right) \approx -\lambda \left( v_i - \frac{1}{n_i} \sum_{j \in U_i} v_j \right), \quad (2.8)$$

approximating the time dependent friction coefficient  $n_i \Lambda$  by a global coefficient  $\lambda$ .

This estimate only approximates  $n_i$  in one of its two instances. Next, we propose an alternative approach. In Section 3.2.2, we test computationally the two assumptions that will be presented, and discuss possible issues.

First, we rewrite Equation 2.7 as follows,

$$\Lambda n_i v_i = \Lambda \sum_{j \in U_i} v_j + \sum_{j=1}^N F_{ij} + P_i F_{p_i}. \quad (2.9)$$

In the absence of net flows in the tissue, we can perform the following approxi-

mation,

$$\sum_{j \in U_i} v_j \approx 0. \quad (2.10)$$

Finally, we take the average number of neighbours to be constant over time, and approximate the friction term by a global friction term  $\lambda$  for all cells and time,

$$\Lambda n_i(t) \approx \lambda > 0. \quad (2.11)$$

**Proposition 1.** *Applying the simplifications presented in this section, the equations of motion become*

$$\begin{aligned} \lambda v_i &= \sum_{j=1}^N F_{ij} + P_i F_{p_i} \\ \frac{dx_i}{dt} &= v_i. \end{aligned} \quad (2.12)$$

## 2.3 Cell differentiation

During animal development, stem cells differentiate and give rise to all the different cell types in the organism. Over time, their protein composition and gene expression profile change irreversibly. The time evolution of a cell can be split in states: we say that a cell transitions to a new state (or cell fate) when their gene expression profile has changed significantly such that it can be considered a different type of cell.

This coarse-grained approach greatly simplifies the analysis: instead of studying the continuous changes in the composition of cells, it is enough to study the discrete dynamics of cell transitions. We will use a unidirectional stochastic process for differentiation, where the transition rates may change depending on the spatial organization of the aggregate.

First, we need some preliminary definitions. Let  $X$  denote an arbitrary cellular state.

**Definition 7.** *The transition rate  $r_{X \rightarrow Y}^{(i)}(t)$  describes the probability per unit time of a cell in state  $X$  to transition to state  $Y$ .*

**Definition 8.** *Let  $\text{state}(i)$  denote the state of cell  $i$ . The state indicator function for cell*

$i$  is

$$\mathbb{1}_X(i) = \begin{cases} 1 & \text{if } \text{state}(i) = X \\ 0 & \text{otherwise.} \end{cases} \quad (2.13)$$

**Definition 9.** *The proportion of cells in state  $X$  is denoted as follows,*

$$\phi_X = \frac{1}{N} \sum_i^N \mathbb{1}_X(i). \quad (2.14)$$

**Remark 4.** *The sum of the proportions is always one, that is,*

$$\phi_A + \phi_B + \phi_C = 1.$$

**Definition 10.** *The proportion of neighbours of cell  $i$  in state  $X$  is denoted as follows,*

$$\langle \mathbb{1}_X(i) \rangle = \frac{1}{n_i} \sum_{j \in U_i} \mathbb{1}_X(j). \quad (2.15)$$

In this work we consider three possible states:  $A$ ,  $B$  and  $C$ . Once the differentiation process begins, cells start transitioning with certain rates from its initial state  $A$  to state  $B$ , and from state  $B$  to state  $C$ . Next, we describe three different expressions for the transition rates. The physical effect of these expressions is compared in Chapter 4.

**Definition 11.** *Let the transition rates be independent of the cell,*

$$r_{X \rightarrow Y}^{(i)}(t) \equiv r_{X \rightarrow Y}(t).$$

*Then, the evolution of the state proportions is described by the following system of differential equations,*

$$\begin{aligned} \dot{\phi}_A &= -r_{A \rightarrow B}\phi_A \\ \dot{\phi}_B &= r_{A \rightarrow B}\phi_A - r_{B \rightarrow C}\phi_B \\ \dot{\phi}_C &= r_{B \rightarrow C}\phi_B, \end{aligned} \quad (2.16)$$

*with initial conditions*

$$\begin{aligned} \phi_A(0) &= a \\ \phi_B(0) &= 1 - a \\ \phi_C(0) &= 0. \end{aligned} \quad (2.17)$$

### 2.3.1 Linear first order differentiation kinetics

The simplest scenario is to consider that cells differentiate with a constant rate, independently of the state of their surrounding cells.

**Definition 12.** *When the transition rates are constant, we denote them by*

$$\begin{aligned} r_{A \rightarrow B} &= p \\ r_{B \rightarrow C} &= q. \end{aligned} \tag{2.18}$$

In this case, Equation 2.16 can be easily solved analytically. The advantage of studying simpler cases first is that they are easier to simulate, and their solution allow us to check if the simulations work accurately.

**Proposition 2.** *When the transition rates are constant, we can compute the analytical solution of Equation 2.16.*

The expression for the analytical solution and the proof of Proposition 2 are detailed in Section A.1 of the Appendix.

However, cells usually communicate through morphogenetic signals which induce cell differentiation. Next, we consider such a case.

### 2.3.2 Nonlinear differentiation kinetics

In this case, transition rates are no longer constant and change over time depending on the surrounding cells. The origin of these nonlinearities is the consequence of feedback in the transition rates.

**Definition 13.** *There is feedback in the transition from one state to another when the transition rate depends on the state of the rest of cells. The feedback factor  $K \geq 0$  determines the strength of this process.*

In particular, we consider the case in which the rates  $r_{A \rightarrow B}$  and  $r_{B \rightarrow C}$  depend on the proportion of cells in state  $A$  following two different approaches. This choice is motivated by experimental evidence in mouse gastruloids (Oriola, Torregrosa-Cortés, et al. [unpublished](#)). For  $K = 0$ , we recover the linear first order kinetics.

Next, we consider a mean field model for feedback. This model assumes that the transition rate of a cell is affected by the global state of the aggregate.

**Definition 14.** *Using mean field feedback, the transition rates depend on the total*

*proportion of cells in state A,*

$$\begin{aligned} r_{A \rightarrow B} &= \frac{p}{1 + K\phi_A} \\ r_{B \rightarrow C} &= \frac{q}{1 + K\phi_A}. \end{aligned} \tag{2.19}$$

Note that, in this case, the transition rates at each time are equal for all cells. Therefore, we can solve Equation 2.16 numerically.

The third approach implies paracrine signalling, that is, cell-cell signalling that induces changes in neighbouring cells. Intuitively, if  $A$  cells are evenly distributed across the aggregate, this will be similar to Equation 2.19 for all cells.

**Definition 15.** *Using cell-cell signalling feedback, the transition rates depend on the proportion neighbours in state A,*

$$\begin{aligned} r_{A \rightarrow B}^{(i)} &= \frac{p}{1 + K\langle \mathbf{1}_A(i) \rangle} \\ r_{B \rightarrow C}^{(i)} &= \frac{q}{1 + K\langle \mathbf{1}_A(i) \rangle}. \end{aligned} \tag{2.20}$$

Usually, cell states are not evenly mixed across the aggregate, so this method mirrors the experimental results more accurately. However, it is not possible to describe a numerically solvable system. The simulation presented in this thesis integrates cell-cell signalling feedback and offers a solution to the evolution of the proportions of each state over time (see Chapter 4).

Making sure that the model simulates the previous differentiation kinetics according to the known solutions will be crucial to confidently implement cell-cell feedback.

## 2.4 Nondimensionalization

Nondimensionalization is a process used to remove the units from variables. It involves scaling the variables in a system by appropriate factors.

**Definition 16.** *The reference values used to scale each variable are*

$$(F_0, R_0, T_0).$$

**Notation 3.** *Dimensionless variables are denoted by a tilde  $\sim$ .*

**Example 1.** Let  $[F]$ ,  $[T]$ ,  $[L]$  be the units of force, time, and length respectively. Then, the units of friction  $\lambda$  are  $\frac{[F][T]}{[L]}$ , and it is nondimensionalized w.r.t. the reference values as

$$\tilde{\lambda} \equiv \lambda \frac{1}{F_0 T_0} \frac{R_0}{T_0}.$$

Following this process, the dimensionless variables of the model are

$$\tilde{F} \equiv \frac{F}{F_0}, \quad \tilde{t} \equiv \frac{t}{T_0}, \quad \tilde{r}_{X \rightarrow Y} \equiv T_0 r_{X \rightarrow Y}, \quad (2.21)$$

and

$$\tilde{x} \equiv \frac{x}{R_0}, \quad \tilde{v} \equiv v \frac{T_0}{R_0}, \quad \tilde{\lambda} \equiv \lambda \frac{R_0}{F_0 T_0}. \quad (2.22)$$

**Proposition 3.** The dimensionless equations of motion are

$$\begin{aligned} \tilde{\lambda} \tilde{v}_i &= \sum_{j=1}^N \tilde{F}_{ij} + P_k \tilde{F}_{p_i} \\ \frac{d\tilde{x}_i}{d\tilde{t}} &= \tilde{v}_i, \end{aligned} \quad (2.23)$$

with initial conditions

$$\tilde{v}_i(0) = \tilde{v}_{0_i}, \quad \tilde{x}_i(0) = \tilde{x}_{0_i}. \quad (2.24)$$

*Proof.* The proof of Proposition 3 is derived in Section A.1 of the Appendix.  $\square$

These are the final equations used to simulate the motion of the cells in the aggregate.

# Chapter 3

## Implementation of the model

All the code developed for this project is available on the [Multiscale Physics of Living Systems Group's GitHub](#)<sup>1</sup> and on my [personal GitHub](#)<sup>2</sup>, and the system specifications are detailed in Section A.2.1.

This chapter outlines the framework used to program the processes described in Chapter 2, details some algorithms and gives computational insight of the system. The model consists of two parts: it simulates the motion of the cells solving Equation 2.23 using numerical methods and reproduces the state transitions as a stochastic process.

### 3.1 Computational framework and numerical methods

#### 3.1.1 Framework

The code has been created using the Julia package for multicellular agent-based modelling [CellBasedModels.jl](#)<sup>3</sup>, developed by Torregrosa-Cortés et al. [unpublished](#), along with the examples provided in the documentation.

A brief description of each program along with the required packages to run them can be found in Section A.2, and a general outline of the code is provided in Program 11.

---

<sup>1</sup><https://github.com/MPoLS-lab/villegas-morral-msc-thesis>

<sup>2</sup><https://github.com/villegas-morral/masters-thesis>

<sup>3</sup><https://github.com/dsb-lab/CellBasedModels.jl>

Note that the computational model works with dimensionless units. Once the simulations are finished, these can be converted to extract and plot the data in physical units.

### 3.1.2 Integration of the equations of motion

The method chosen to solve the equations of motion for each cell is Heun's method, which can be derived from Euler's method (Süli and Mayers 2003; Bronson 2022). Consider the following differential equation,

$$\frac{dx}{dt} = v(t, x).$$

Given a timestep  $\Delta t$  and a number of steps  $l$ , its solution is approximated by the discrete function  $\{x_n\}_{n=0,\dots,l}$  defined over  $\{t_n\}_{n=0,\dots,l}$  such that

$$x_n \approx x(t_n), \quad t_{n1} - t_n = \Delta t.$$

The values  $x_0, v_0$  at time  $t_0$  follow from the initial condition.

In the forward Euler method, the rest of the values are approximated by taking

$$x_{n+1} = x_n + v(t_n, x_n)\Delta t, \quad (3.1)$$

while in the backward Euler method, they follow from

$$x_n = x_{n+1} - v(t_{n+1}, x_{n+1})\Delta t. \quad (3.2)$$

Heun's method is a simple predictor-corrector method which combines these two:

1. Compute  $x_n$  and  $v(t_n, x_n)$ .
2. Compute an initial prediction  $x_{n+1,p}$  using the forward Euler method,

$$x_{n+1,p} = x_n + v(t_n, x_n)\Delta t.$$

3. Isolate  $x_{n+1}$  in Equation 3.2 and compute the corrector  $x_{n+1,c}$  according to the backward Euler method using the predictor  $x_{n+1,p}$  obtained in the previous step,

$$x_{n+1,c} = x_n + v(t_{n+1}, x_{n+1,p})\Delta t.$$

4. The final approximation  $x_{n+1}$  is the average,

$$x_{n+1} = x_n + \frac{\Delta t}{2} (v(t_n, x_n) + v(t_{n+1}, x_{n+1,p})) . \quad (3.3)$$

### 3.1.3 Algorithm for cell differentiation

Given probabilities  $p, q$  as described in Definition 12, cell differentiation for an arbitrary cell  $i$  follows the next rules for each timestep  $\Delta t$ .

1. If the evolution of states is turned off or  $\text{state}(i) = C$ , stop.
2. Count the proportion of neighbours in state  $A$ , that is,  $\langle \mathbb{1}_A(i) \rangle$ .
3. If  $\text{state}(i) = A$ 
  - i. Compute  $r_{A \rightarrow B}$ .
  - ii. Choose  $\xi \rightarrow U(0, 1)$ .
  - iii. If  $\xi < r_{A \rightarrow B} \Delta t$ , then cell  $i$  transitions and  $\text{state}(i) = B$ .
4. If  $\text{state}(i) = B$ 
  - i. Compute  $r_{B \rightarrow C}$ .
  - ii. Choose  $\xi \rightarrow U(0, 1)$ .
  - iii. If  $\xi < r_{B \rightarrow C} \Delta t$ , then cell  $i$  transitions and  $\text{state}(i) = C$ .

## 3.2 Numerical validation of the global friction approximation

This section outlines the procedures followed to obtain a functional and efficient computational model, focusing on the validation of the friction approximation stated in Equation 2.11.

### 3.2.1 Aggregate stability

Some parameters can induce numerical errors if not tuned appropriately. For example, when a cell in the outer layer divides, the cells that were in contact with the mother cell may detach from the aggregate due to numerical errors (see Figure 3.1). In order to obtain a stable aggregate, we will adjust the timestep ( $\Delta t$ ), the overlap parameter ( $\alpha_{ov}$ ), and the friction coefficient ( $\lambda$ ).

The stability of a parameter can be tested using Program 12, which implements a function that returns whether the aggregate consists of a single connected component.

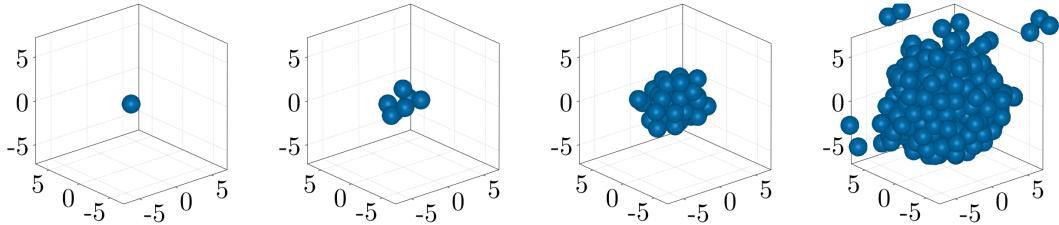


Figure 3.1: Unstable simulation.

Complex models often require smaller timesteps, which increase computational costs. Balancing this trade-off between model complexity and computational efficiency is essential for creating an effective model.

During cell division, two daughter cells emerge from a mother cell. The overlap factor determines how closely these daughter cells are placed in the aggregate after the mother cell is removed. This simplification accounts for the repulsion force between these cells before they completely split, so that we can modify  $\alpha_{ov}$  freely.

Since we are simulating *in vitro* cell cultures, the physical parameters could be in principle measured. In our study, friction accounts for the binding and unbinding connecting neighbouring cells.

Next, we focus on the stability of the aggregate formation, turning off differentiation. Since cell rearrangements are not relevant, we simplify the computations by setting active protrusions to zero ( $F_{pi} = 0$ ).

### 3.2.2 Global friction and average number of neighbours

The aim of this section is to analyse the average number of neighbours in the system in order to identify and correct potential issues, and to compute an approximation for it. This leads to the expression selected for  $F_{ij}$  in Equation 2.3.

Let  $N(t)$  denote the total number of agents in the system, and  $\langle n_i \rangle(t)$  denote the average the number of neighbours of a cell over time,

$$\langle n_i \rangle(t) = \frac{1}{N(t)} \sum_{i=1}^N n_i(t). \quad (3.4)$$

We consider a simpler (see Equation 2.3) force profile for  $F_{ij}$  that will be polished after the analysis,

$$F_{ij} = \begin{cases} \frac{(x_i - x_j)}{d_{ij}} F_0 f_r(d_{ij}) \alpha_{\text{adh}} & \text{if } d_{ij} < 2\mu r \\ 0 & \text{otherwise.} \end{cases} \quad (3.5)$$

The final equations of motion presented in Chapter 2 assume the simplification

$$\Lambda n_i(t) \equiv \lambda > 0, \quad (2.11)$$

presented in Section 2.2.3. The value for  $\lambda$  will be chosen using  $\langle n_i \rangle$ .

Therefore, in this section we consider the following equations of motion,

$$\begin{aligned} \Lambda n_i v_i &= \sum_{j=1}^N F_{ij} \\ \frac{dx_i}{dt} &= v_i. \end{aligned} \quad (3.6)$$

Note that we are not addressing the absence of net flows yet, presented in Equation 2.10 before the global friction. For convenience, this is discussed in the next section (3.2.3).

Since the computational implementation uses the dimensionless form of the equations, the typical value of the friction coefficient  $\tilde{\Lambda} n_i$  is expected to be of the order of 1. The number of neighbours should be of the order of 10, so we set  $\tilde{\Lambda} = 0.1$ .

### Analysis of $\langle n_i \rangle$ over time

During proliferation, the average number of neighbours is expected to stabilize over time. In particular, since cells in the centre of the aggregate have the maximum number of neighbours, once the distance to the outer layer is larger than the neighbour range, this number should not change.

In CellBasedModels.jl, aggregates are stored in objects which allow access to past timestamps. Using Program 4, we grow five aggregates up to 1000 agents (see Figure 3.2), and compute  $\langle n_i \rangle$  against  $N(t)$ . This plot is displayed in Figure 3.3.

This data shows that, in this simulation, the average number of neighbours in-

creases with  $N$  and could be roughly approximated to

$$\langle n_i \rangle(t) \approx 0.1N(t). \quad (3.7)$$

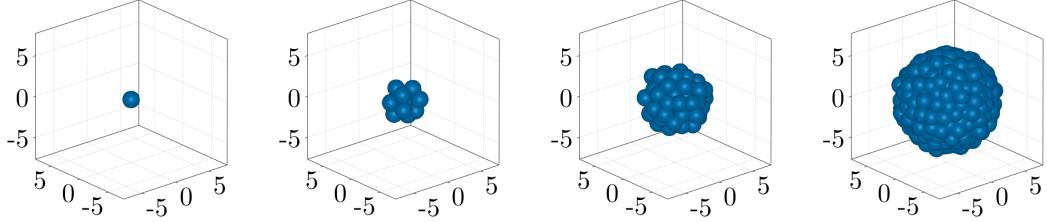


Figure 3.2: Proliferation of an aggregate using Equations 3.5 and 3.6.

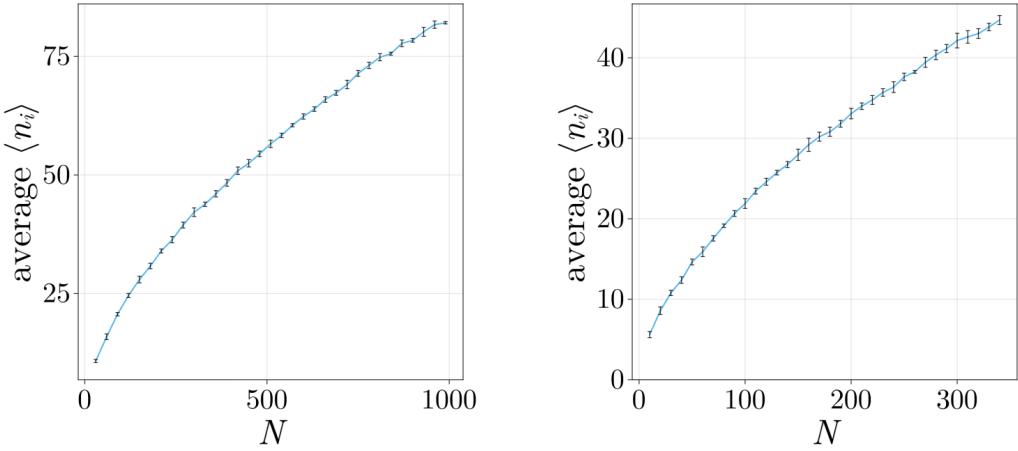


Figure 3.3:  $\langle n_i \rangle$  against  $N(t)$ . Averaged over five realizations. Bars indicate the standard deviation of averaging the five systems.

However, this is not what one would expect from a closely packed structure. We continue the analysis and, later in the section, propose a solution.

Henceforth, we will use Program 5 to focus on a single realization due to the small standard deviation observed. First, we plot the number of neighbours (see Figure 3.4).

The lowest number of neighbours, corresponding to the outer layer cells, remains constant as expected. However, it is the cells with largest  $n_i$  that experience an increase in neighbours, that is, the inner cells. To visualize the issue, we plot the neighbours of the cell with the highest  $n_i$  in the final timestamp alongside the aggregate itself in Figure 3.5.

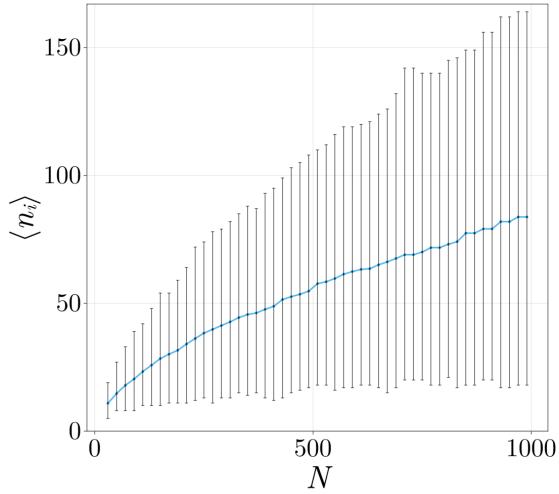


Figure 3.4: Average number of neighbours against  $N(t)$ . Bars indicate the highest and lowest number of neighbours.

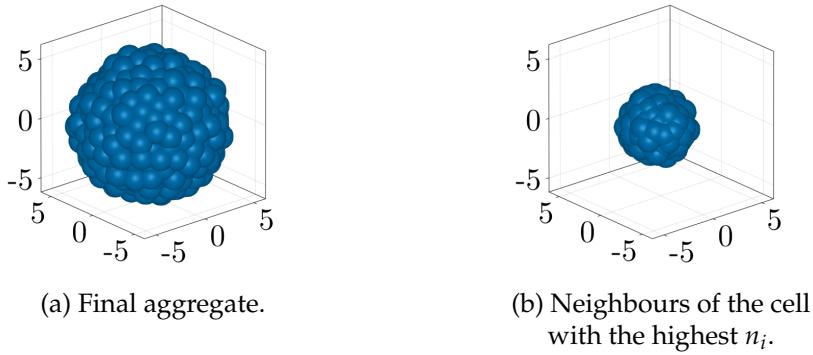


Figure 3.5: Aggregate of  $N = 900$  and set of  $n_i = 151$  neighbours.

The issue is that cells in the inner part of the aggregate become trapped in a dense configuration with excessive overlap, which is physically implausible. This is a common drawback when using centre-based models (Van Liedekerke et al. 2015).

### Proposed solution

A possible way of addressing this issue can be using a different method to count neighbours, such as considering Gabriel neighbours (Oriola, Marin-Riera, et al. 2022). If we used this method, we should either approximate  $\lambda$  by Equation 3.7 depending on the magnitude of the aggregates to simulate, or use the model with variable friction  $n_i\Lambda$ .

Decreasing the proliferation rate to allow cells to relax and accommodate with

less overlap does not resolve the problem either. To ensure that cells are able to reorganize during the differentiation process, we instead propose a change in the force profile to prevent excessive overlap.

In the definition for  $F_{ij}$  stated in Equation 3.5, the expression for the attractive and the repulsive regimes described in Remark 2 is the same. Still, repulsion forces are not enough to separate the cells overlapped in the centre. We thus consider a force profile that splits the two regimes and controls the difference in strength through a repulsion factor,

$$F_{ij} = \begin{cases} \frac{(x_i - x_j)}{d_{ij}} F_0 f_r(d_{ij}) \alpha_{\text{adh}} \alpha_{\text{rep}} & \text{if } d_{ij} \leq 2r \\ \frac{(x_i - x_j)}{d_{ij}} F_0 f_r(d_{ij}) \alpha_{\text{adh}} & \text{if } 2r < d_{ij} < 2\mu r \\ 0 & \text{otherwise,} \end{cases} \quad (2.3)$$

Next, we look for an appropriate repulsion factor using Program 6. Recall that, for  $\alpha_{\text{rep}} = 1$ , we recover the force expression from Equation 3.5.

Our final aim is to reliably simulate aggregates of the magnitude of 500 cells, so we consider a system of  $N = 800$ . Varying  $\alpha_{\text{rep}} \in \{1, 2, 2.5, 3\}$ , the cells with the highest  $n_i$  have, respectively, 142, 45, 30 and 25 neighbours. These neighbourhoods and the aggregates they belong to are shown in Figure 3.6.

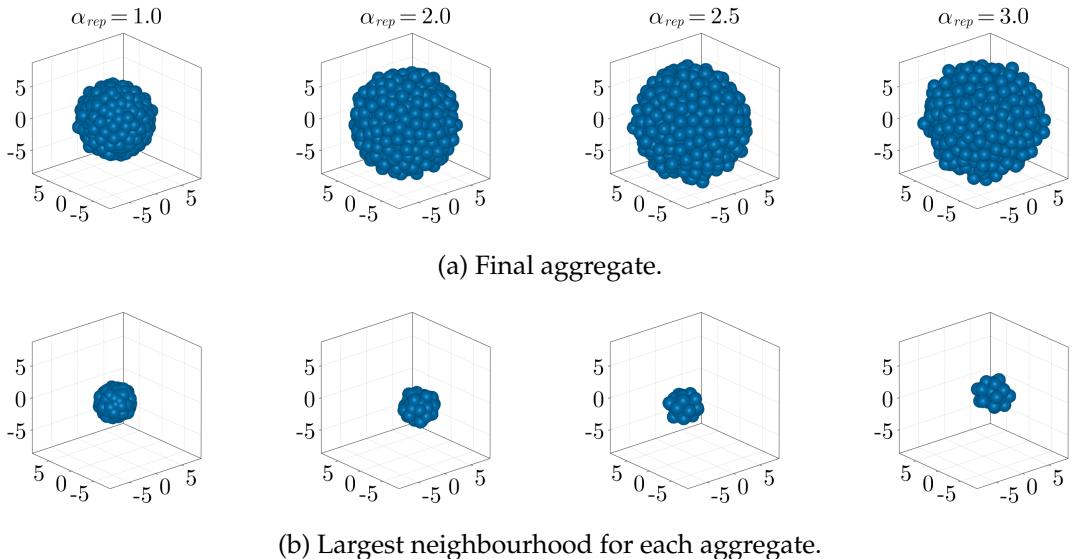


Figure 3.6: Aggregates and sets of neighbours by varying  $\alpha_{\text{rep}}$ .

Figure 3.8 illustrates the average number of neighbours in terms of the number

of cells. Values  $\alpha_{\text{rep}} \in \{2.5, 3\}$  yield the best results. The averages of neighbours are approximated as their average over time,

$$\begin{aligned}\langle n_i \rangle(t)_{2.5} &\approx 8 \pm 5, \\ \langle n_i \rangle(t)_{3.0} &\approx 7 \pm 4\end{aligned}\tag{3.8}$$

We determine to set  $\alpha_{\text{rep}} = 2.5$  when simplifying the model using global friction (see Program 8), given that it is the lowest factor that offers a realistic output, and approximate  $\langle n_i \rangle(t) \approx 10$ , thus considering the friction coefficient

$$\tilde{\lambda} = 10\tilde{\Lambda} = 0.1.\tag{3.9}$$

To further discourage the possibility that cells cluster because of a numerical problem during development, we tried setting  $\alpha_{\text{rep}} = 1$  in an aggregate grown using  $\alpha_{\text{rep}} = 2.5$ . The aggregate rapidly compressed (see Figure 3.7), proving that  $\alpha_{\text{rep}} = \alpha_{\text{atr}} = 1$  are not feasible in this model.

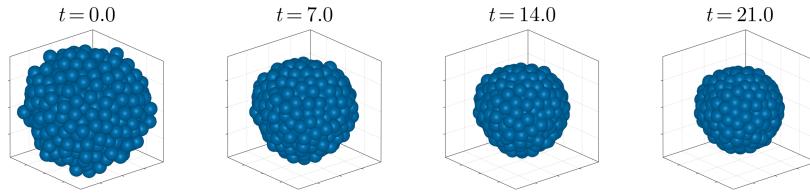


Figure 3.7: Changing  $\alpha_{\text{rep}}$  from 2.5 to 1.

Finally, the protrusion force is also split into repulsive and attractive regimes,

$$F_{p_i} = \begin{cases} -\frac{(x_i - x_j)}{d_{ij}} D \alpha_{\text{rep}} & \text{if } d_{ij} \leq 2r \\ -\frac{(x_i - x_j)}{d_{ij}} D & \text{if } 2r < d_{ij} < 2\mu r. \end{cases}\tag{2.3}$$

The appropriate implementation of this factor is addressed in Section 4.1.3.

### 3.2.3 Absence of net flows

This section tests using different approaches the assumption that net flows are negligible,

$$\sum_{j \in U_i} v_j \approx 0.\tag{2.10}$$

First, we simulate the system using the equations of motion before the approximation,

$$\begin{aligned}\Lambda n_i v_i &= \Lambda \sum_{j \in U_i} v_j + \sum_{j=1}^N F_{ij} + P_i F_{p_i} \\ \frac{dx_i}{dt} &= v_i,\end{aligned}\tag{3.10}$$

However, when including the sum of velocities term, our code turns out to be numerically unstable (see Program 9).

The simplification consists of removing the weight of the sum of velocities from the equations of motion, but it does not actually set this term to be zero. Therefore, testing Hypothesis 2.10 with the model with variable friction proposed in Program 6 is still significant. Program 10 verifies that the expression is indeed null.

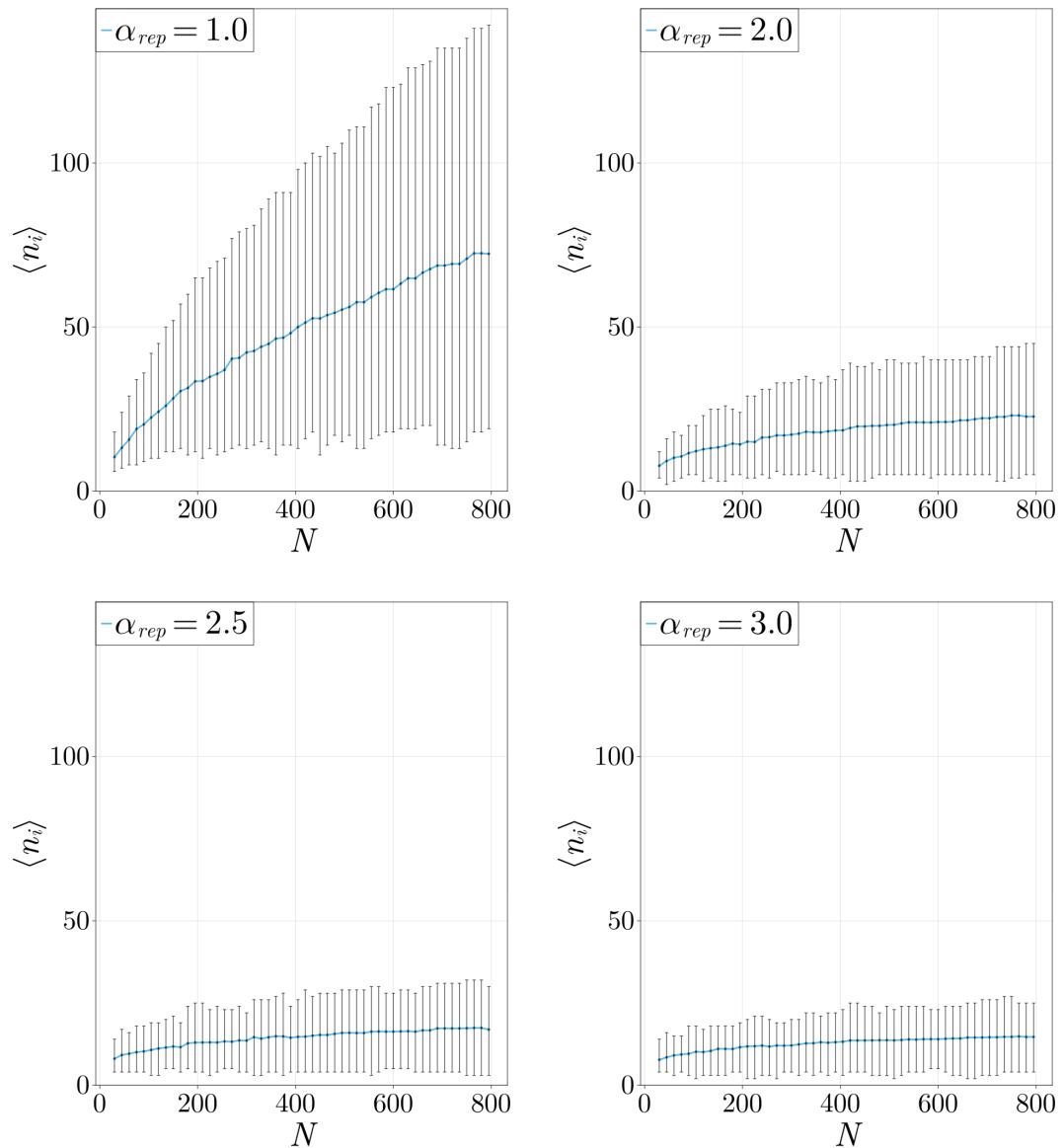


Figure 3.8: Comparison of  $\langle n_i \rangle$  for different values of  $\alpha_{\text{rep}}$ . Bars indicate the minimum and maximum number of neighbours for each  $N$ .

# Chapter 4

## Test examples and applications

In this chapter, we display the performance of the model by tuning it to study the early stages of mouse gastruloid formation. Mouse gastruloids are organoids aggregated from mouse embryonic stem cells (mESCs). The experiments presented in Oriola, Torregrosa-Cortés, et al. [unpublished](#) show that:

- Cell-cell communication influences cell differentiation and the timing of symmetry breaking in gastruloids
- T dynamics is highly dependent on the initial T+ fraction.

We implement the model using the feedback expression proposed in the aforementioned study. Our aim is to analyse using the model the effect of cell-cell communication in the differentiation process, showing how spatial organization affects the evolution of stem cells, and to offer a 3D visualization for the early stages of gastruloid formation. A realization of the differentiation process is shown in Figure 4.1.

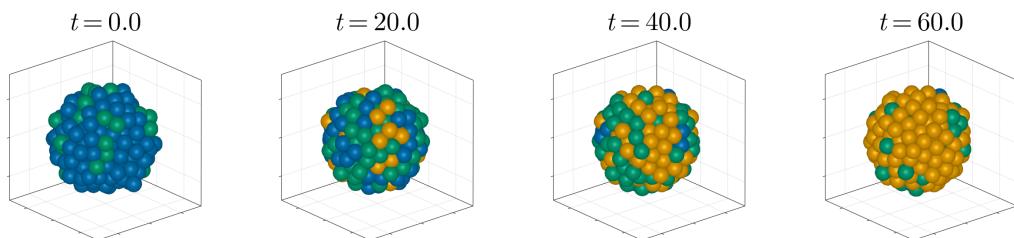


Figure 4.1: Simulation of the differentiation process using cell-cell feedback.  
Time in hours.

First, let us provide a description of the experimental model simulated, along with some insight on the functionality of the simulation.

## 4.1 Physical description

### 4.1.1 Experimental system

The simulation is used to recreate the expression of Brachyury (T), a gene that plays an essential role in the formation of embryonic mesoderm. States T- and T+ denote whether the cells express the gene, and are distinguished based on GFP (green fluorescent protein) expression. We do not consider apoptosis (cellular death), as it is not significant in gastruloid formation.

The different stages during early differentiation can be split into three unidirectional states. Let state *A* represent T-, state *B* represent T+, and *C* represent further cellular fates. The experiment consists in creating cell aggregates mixing T+ cells and T- cells under controlled proportions and studying their development. We simulate this by growing an aggregate of cells in state *A* from a single cell, which resembles the final structure observed in the cultures aggregated experimentally. Then, we initialize a fixed proportion of cells in state *B* chosen at random.

The parameters shown in the [Nomenclature](#) are chosen to mirror the experimental data, taken from the following sources.

- Cell radius is taken from Pillarisetti et al. [2009](#).
- Average division range is taken from Roccio et al. [2013](#).
- Neighbouring and force ranges are taken from Saiz et al. [2020](#).
- Differentiation parameters are taken from Oriola, Torregrosa-Cortés, et al. [unpublished](#).

Unless stated otherwise, initial aggregates are grown up to 300 cells, following Brink et al. [2014](#). Next, we compute some measurements to serve as a reference for the rest of the study.

### 4.1.2 Proliferation speed

In our model, proliferation appears in two contexts: when the initial aggregate is formed and when it occurs during differentiation. In the former, the average division time  $\tau_{\text{div}}$  can be chosen at will as long as the simulation is numerically stable, since the aim is to get the structure. During differentiation, however, the parameter has to be physically accurate in order to tie differentiation dynamics and tissue mechanics.

In the computation of the initial aggregate, we start from a single cell and grow it up to a given number of cells, whereas in differentiation we start from a system and grow it for a given time. Proliferation requires allocating space before starting the process. Let  $N_0$  be the initial number of cells, we expected the number of cells  $N_e(t_f)$  after time  $t_f$  to increase exponentially,

$$N_e(t_f) = N_0 2^{\frac{t_f}{\tau_{\text{div}}}}. \quad (4.1)$$

Due to the stochastic nature of proliferation, the simulation sometimes results in larger values, and so we preallocate  $1.4N_e(t_f)$  cells. In Figure 4.2 depicts the growth functions for the values of  $\tau_{\text{div}}$  used in each process (see Program 14).

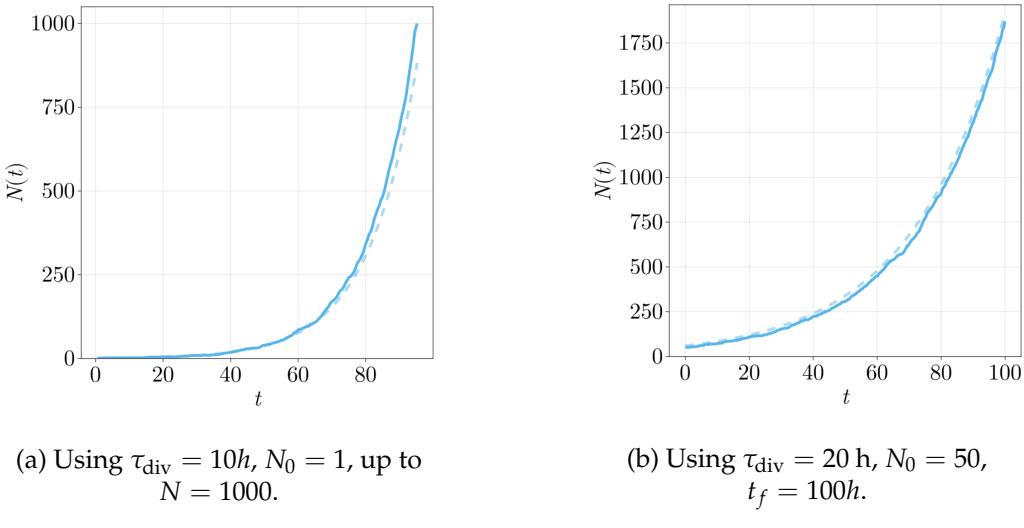


Figure 4.2: Growth of the aggregate in the simulation.  
Dashed lines correspond to  $1.4N_e(t_f)$ .

### 4.1.3 Protrusion force

The effect of protrusions is adjusted tuning the following parameters,

$$\rho = (\tilde{D}, \tilde{k}_{p_{\text{on}}}, \tilde{k}_{p_{\text{off}}}).$$

The amount of movement due to protrusions can be adjusted increasing  $D$  and  $k_{p_{\text{on}}}$ . Greater force strength attracts the cells faster while they are connected, and greater activation rate increases the amount of cells connected to another at each time. In opposition, when protrusions and the growth are disabled, the aggregate tends towards a stable configuration and cells eventually stop moving.

We use the protrusion strength  $D$  presented in Yusko and Asbury 2014; Moore, Roca-Cusachs, and Sheetz 2010. Still, we do not have experimental data for the protrusion activation and deactivation rates. In order to model the random movement of cells seen experimentally using protrusion forces, we set the parameters to

$$\rho_1 = (\tilde{D}, \tilde{k}_{p_{\text{on}}}, \tilde{k}_{p_{\text{off}}}) = (10, 20, 10). \quad (4.2)$$

When the protrusions follow this values, the aggregate slightly contracts, but the number of neighbours remains numerically stable. This happens because cells in the outer layer are connected to cells on the inside.

If we want to increase the amount of cell movement, increasing  $\tilde{D}$  does not suffice: it yields densely packed aggregates as those presented in Section 3.2.2. This can be corrected decreasing the protrusion activation rate. We propose the following numerically stable tuple to generate faster cell displacements,

$$\rho_2 = (\tilde{D}, \tilde{k}_{p_{\text{on}}}, \tilde{k}_{p_{\text{off}}}) = (50, 2, 0.5). \quad (4.3)$$

Despite the value for the force does not correspond to experimental value, we argue that we can use the tuple as a whole because it causes the sought effect.

#### 4.1.4 Measure of movement

In this section, we present a method for measuring movement aggregates and apply it to analyse effect of the protrusion force.

When the number of cells  $N$  is constant between times  $t_0$  and  $t_n$ , we can use the ensemble average approximation to the mean squared displacement (MSD) to measure the displacement of cells inside the aggregate (Rosen, Grant, and Dallon 2021). Let  $J$  be the number of saved timestamps and  $t = \{t_n\}_{n=0,\dots,J}$  be the associated simulation time.

**Definition 17.** *Let  $t_0$  be the initial simulation time. The MSD at time  $t_n$  is defined as follows,*

$$d_2(t_n) = \frac{1}{N} \sum_{i=1}^N |x_i(t_n) - x_i(t_0)|^2. \quad (4.4)$$

Let us also define a function to measure the movement in a chosen interval.

**Definition 18.** *The mean square displacement between times  $t_1$  and  $t_2$  is defined as*

follows,

$$d_2(t_1, t_2) = \frac{1}{N} \sum_{i=1}^N |x_i(t_2) - x_i(t_1)|^2, \quad (4.5)$$

and the MSD step at time  $t_n, n > 0$  as follows,

$$g(t_n) = d_2(t_{n-1}, t_n) \quad (4.6)$$

Using Program 15, we carry out measurements for different cases. For reference, cells have a radius of 5 pm, and the maximum distance between cells in 300-cell aggregates is around 65 pm.

When growth stops, cell move towards a stable configuration, and when active forces are neglected ( $D = 0$ ), cells gradually stop moving. We plot the first 120 h of this process and the MSD step at each timestamp in Figure 4.3.

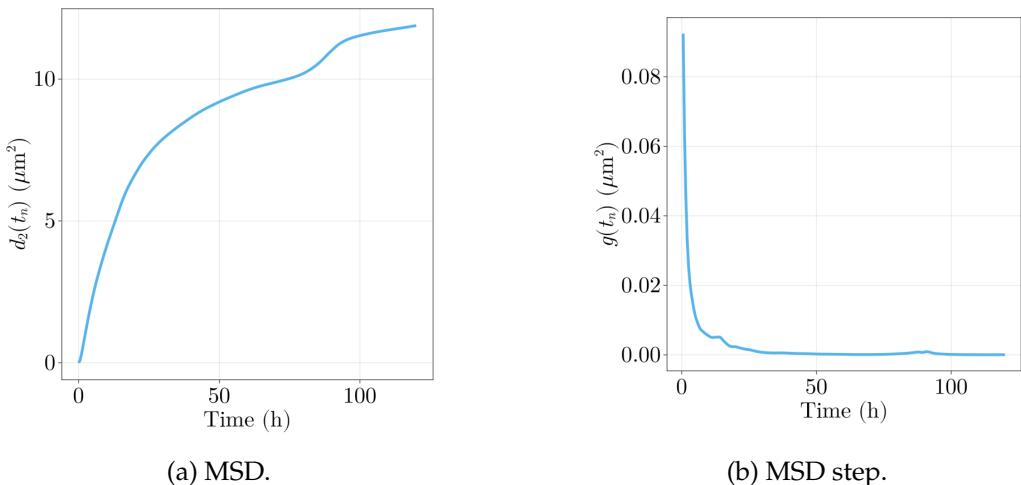
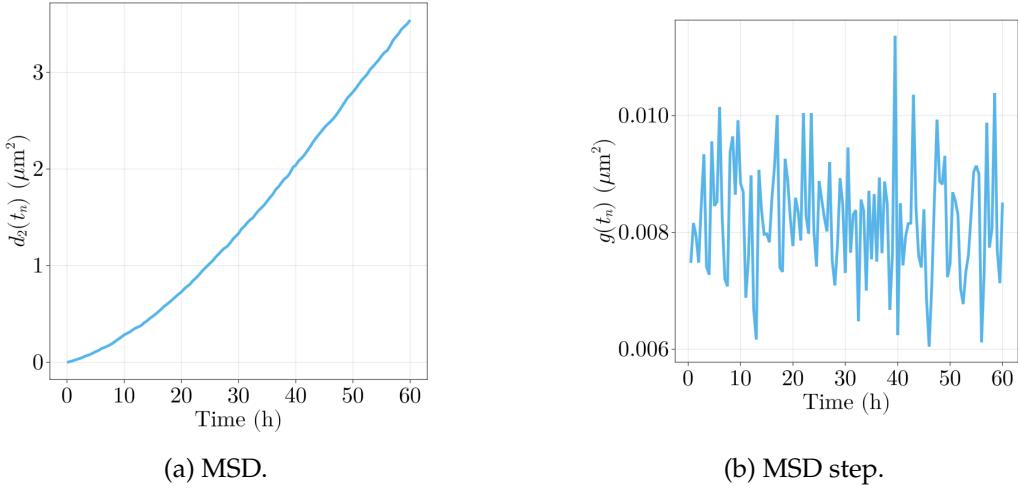


Figure 4.3: Measure of cell movement for  $D = 0$ .

We observe that the MSD reaches a plateau at around  $t = 60$  h. To analyse the effect of protrusions on movement, we initialize aggregates and let them accommodate for this time to obtain more significant data.

Figure 4.4 shows the effect of the protrusion configuration  $\rho_2$ . The MSD shows linear growth, which indicates diffusive behaviour, as expected. The MSD step shows the oscillations caused by the activation and deactivation of the protrusions, and by the cells moving out of the force range of their pairs. Repeating the computations for  $\rho_1$  shows analogous results with smaller values.

Figure 4.4: Measure of cell movement for  $\rho_2$ .

#### 4.1.5 Stable timesteps

The timesteps used for each type of evolution process are shown in Table 4.1. Evolution with proliferation is more unstable and requires a lower timestep, so we will disable this option if possible.

Process	Timestep
Growth of the initial aggregate	0.002
Evolution without proliferation, protrusions disabled	0.002
Evolution with proliferation, protrusions enabled	0.001
Evolution with proliferation, low protrusion strength	0.0005
Evolution with proliferation, high protrusion strength	0.0001

Table 4.1: Timestep values for each evolution process.

## 4.2 Effect of differentiation kinetics on the evolution of the fate proportions

The results reproduced here were obtained using Program 16.

Next, we simulate the three expressions for the transition rates presented in Section 2.3, referred to as *constant rates* ( $p, q$ ), *mean field feedback* ( $r_{A \rightarrow B}, r_{C \rightarrow C}$ ),

and *cell-cell feedback* ( $r_{A \rightarrow B}^{(i)}, r_{B \rightarrow C}^{(i)}$ ),

$$\begin{aligned} r_{A \rightarrow B} &= \frac{p}{1 + K\phi_A}, & r_{B \rightarrow C} &= \frac{q}{1 + K\phi_A} \\ r_{A \rightarrow B}^{(i)} &= \frac{p}{1 + K\langle \mathbb{1}_A(i) \rangle}, & r_{B \rightarrow C}^{(i)} &= \frac{q}{1 + K\langle \mathbb{1}_A(i) \rangle}. \end{aligned} \quad (4.7)$$

The behaviour of cells is stochastic and the changes we want to observe may be minor, so the results obtained from a realization are not enough to determine the accuracy of the model; a low number of cells can also affect this precision (see Figure 4.5). Thus, we average several realizations to obtain more significant results.

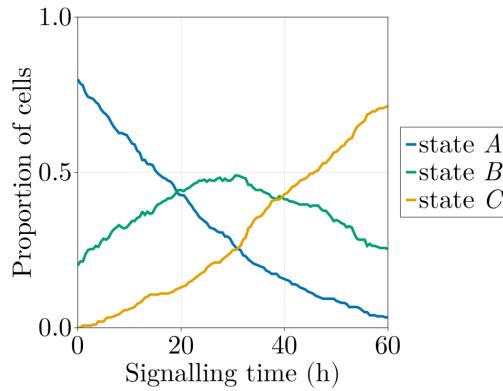


Figure 4.5: Proportions over time for a single realization using cell-cell feedback.

### 4.2.1 Differentiation without proliferation

We can solve the evolution equations for constant rates and for mean field feedback (see Figure 4.6). We compute our results for these cases to assert that the differentiation is properly implemented.

We create an aggregate, stop its growth, and evolve it repeatedly for a time of  $t_e = 60h$ . The initial portion of cells in state  $B$  is set to  $b = 0.2$ , and the protrusion is turned off.

The model replicates the behaviour of the solutions using constant rates and mean field feedback, as shown in Figure 4.7.

Once the model is verified, we simulate the nonlinear differentiation kinetics with cell-cell signalling feedback. This effect has only been observed experimentally and the evolution of the proportions cannot be solved analytically, as

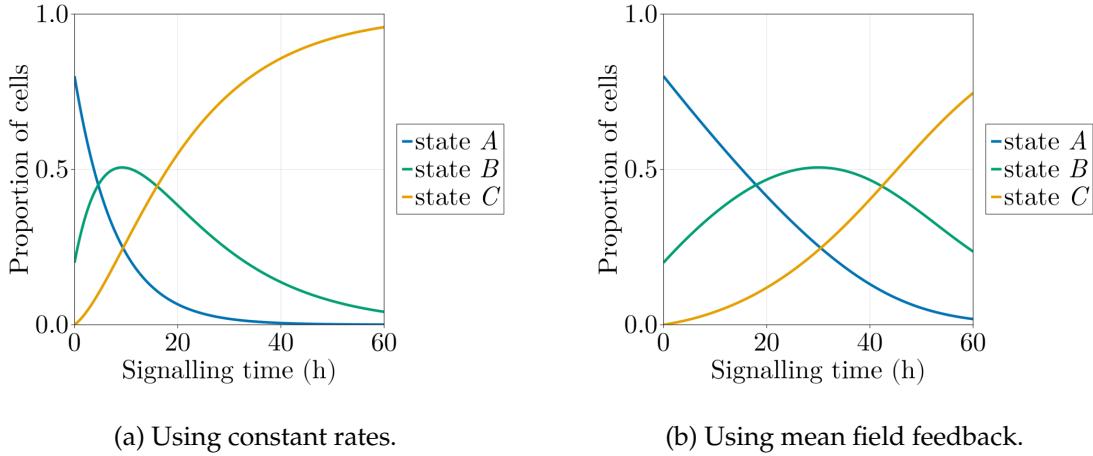


Figure 4.6: Proportions over time computed using the known solutions.

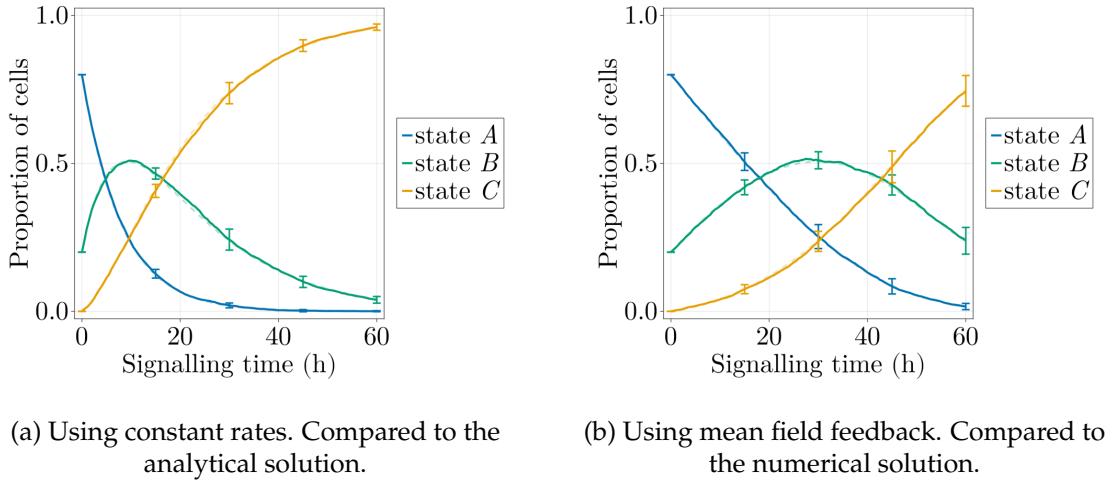


Figure 4.7: Proportions over time computed using the simulation. Averaged over 15 realizations. Dashed lines correspond to the solutions in Figure 4.6, and bars indicate the standard deviation.

it accounts for spatial distribution.

In the first two cases, spatial organization did not affect the differentiation process, so we set  $D = 0$ . However, cell rearrangements might influence the evolution when using cell-cell signalling feedback. We present the results obtained for  $D = 0$  and  $\rho = \rho_1$  in Figure 4.8.

Both situations showcase differences between the cell-cell signalling and the mean field approach: they begin unaltered, and at time  $t = 20$  h state B begins decreasing slightly faster, and state C increasing slightly faster. However, cell protrusions do not seem to have an impact on cell proportions.

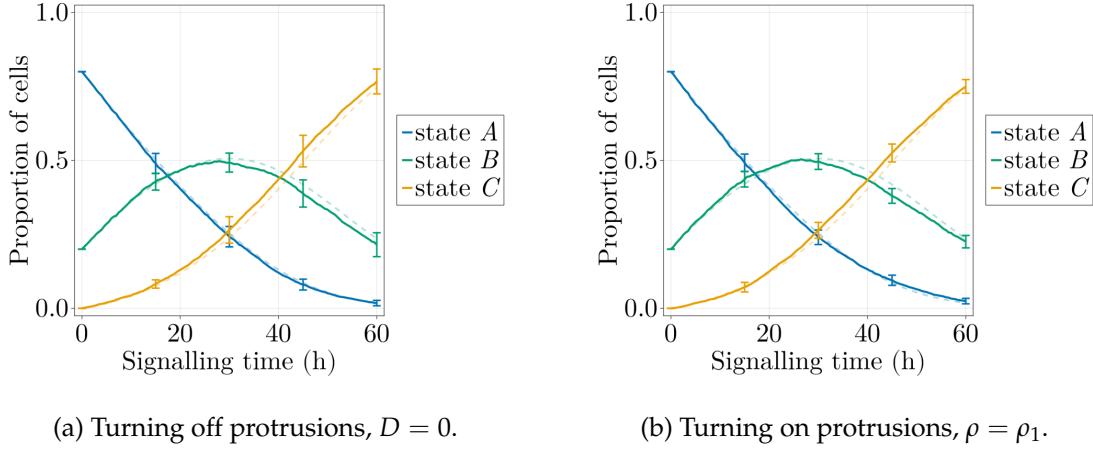


Figure 4.8: Proportions over time computed using the simulation. Averaged over 15 realizations. Dashed lines correspond to the mean field solution, and bars indicate the standard deviation.

From now on, we focus on the cell-cell signalling case.

### 4.2.2 Differentiation with proliferation

The results reproduced here were obtained using Program 17.

In the previous experiment we disable proliferation during differentiation, given that implementing both processes at the same time is less numerically stable. Still, this process may affect the states' evolution. We perform again the computations to ensure that, under these conditions, we can turn it off to study the aggregate. The timestep is reduced, the initial population is 60 and the average division time is set to  $\tau_{\text{div}} = 20$  h.

We perform a simulation using mean field feedback with proliferation and determine that, as expected, the aggregate still follows the solution. Using this as a reference, we simulate two aggregates varying the protrusion strength. The plots are shown in Figure 4.9.

The stochasticity of growth makes these models very variable. Still, it seems that the states' evolution follows the same form as when cells do not proliferate. This behaviour is expected, since the rate of proliferation of all cells is the same. Using this result, we argue that, under the aforementioned conditions, we can study the differentiation process turning off differentiation.

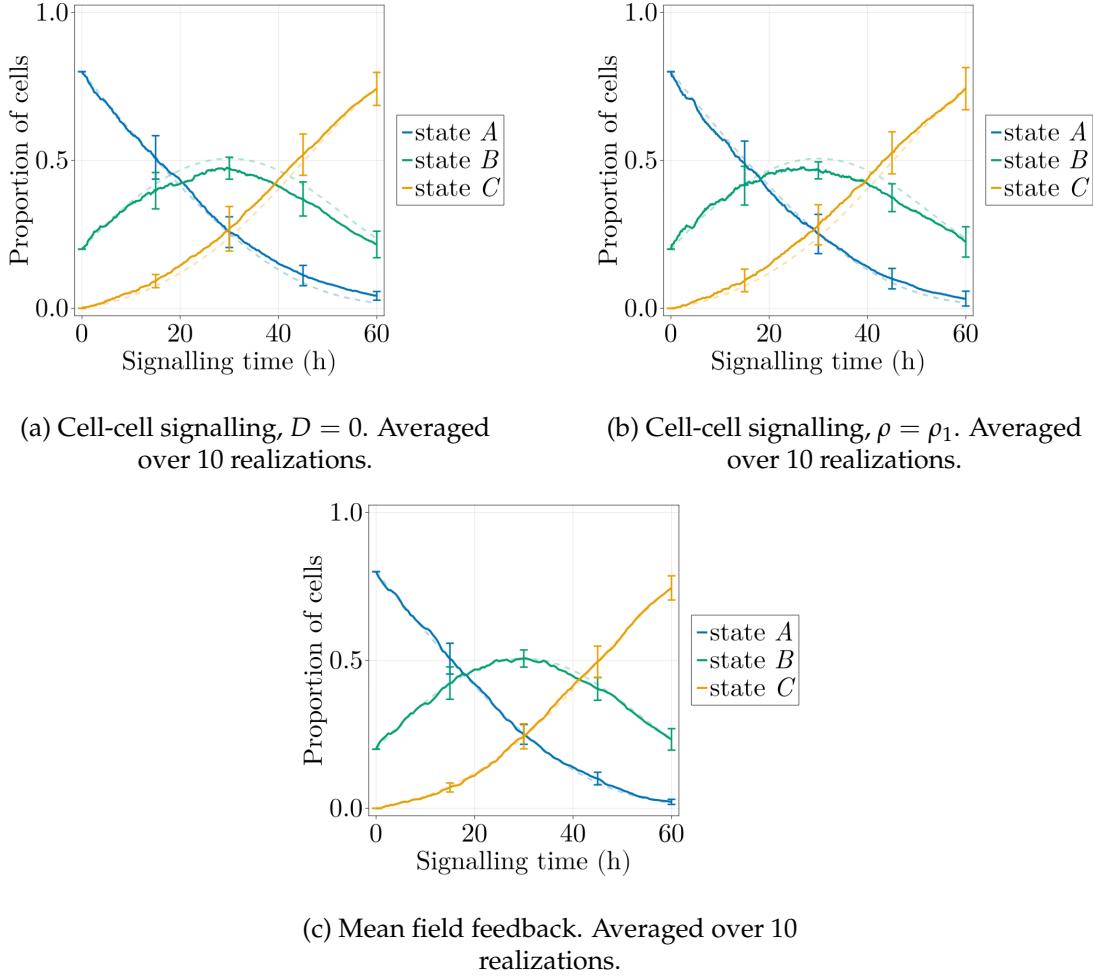


Figure 4.9: Proportions over time with proliferation, compared to the mean field solution (dashed lines).

### 4.3 Proportion of $B$ cells in terms of $b$

The results reproduced here were obtained using Program 18.

In this section, we compute the evolution of the proportion of cells in state  $B$  at a fixed time  $t_n$ ,  $\phi_B(t_n)$ , according to its initial proportion,  $\phi_B(0) = b$ . The objective of this approach is to analyse how initial cell population affects cellular differentiation and get another insight on the effect of feedback.

We will compute the function for times  $t \in \{0, 15, 30, 45, 60\}$  (h). The plots for the constant and mean field rates are shown in Figure 4.10 to serve as a reference for the rest of the section. The analogous plots for states  $A$  and  $C$  in terms of  $b$  are displayed in Section A.3 of the Appendix for the sake of completion (see Figure 5.1).

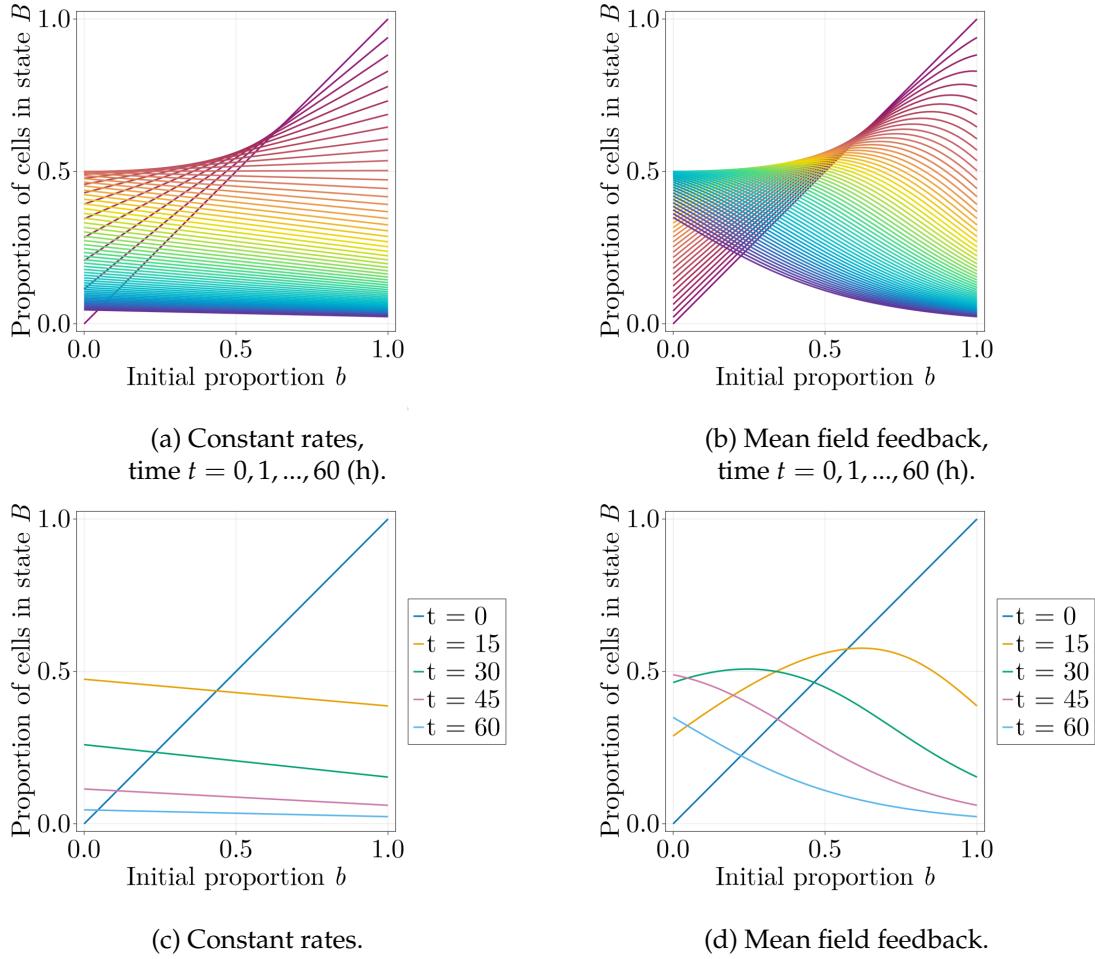


Figure 4.10:  $\phi_B(t_n)$  against  $b$  using the known solutions.

Let us analyse the solutions. For constant rates (or linear differentiation kinetics) the time plots rapidly become decreasing, and in Figure 4.10a the proportion of  $B$  cells is always larger at  $t = 15$  h. This simple model can reproduce a maximum in the  $B$  population (see Figure 4.6), still, it cannot reproduce the nonlinear dependence observed experimentally for this plot (Oriola, Torregrosa-Cortés, et al. [unpublished](#)).

For mean field feedback (a type of nonlinear differentiation kinetics), the plots for approximately the first half of the simulation exhibit a maximum, whereas the other half accomplishes its maximum proportion when  $b = 0$ . Whenever the time plot is above than the identity line ( $t = 0$ ), the proportion of cells at that time is greater than the corresponding  $b$ . Unlike in the linear case, the

proportion of  $B$  cells increases and decreases over time, namely

$$b = 0 : \quad \phi_B(45) > \phi_B(30) > \phi_B(60) > \phi_B(15) > b,$$

$$b = 0.25 : \quad \phi_B(30) > \phi_B(15) > \phi_B(45) > b > \phi_B(60).$$

For small  $b$ , the proportion at later times is much greater than the linear case, since more cells are in state  $A$  and thus inhibit the differentiation. As  $b$  approaches 1, the plot resembles the linear case.

### 4.3.1 Using cell-cell feedback

In order to visualize the effect of feedback and protrusion force, we compute simulations using cell-cell feedback model varying  $\rho$  (see Figure 4.11).

Obtaining one of the averages shown in Figure 4.11 requires the simulation of 70 differentiation process, since we compute them for seven different  $b$ . Nevertheless, the plots to the left, which represent a realization with five  $b$  points, reflect the averages. We conclude that the general behaviour of the differentiation kinetics is accurately described using this simpler method.

Regarding random cell displacements, setting  $\rho = \rho_2$  does not produce visible effects with respect to  $D = 0$ , and  $\rho = \rho_1$  displays a slightly different behaviour at some points (see  $b = 0$ ). The expected behaviour is that, for high noise, the aggregate resembles homogeneity and outputs similar results to the mean field model. However, the data obtained under the aforementioned simulation conditions does not suffice to affirm such a thing.

We further illustrate case  $\rho = \rho_1$  by computing the proportion of cells in state  $A$  and  $C$  compared to their respective solutions (see Figure 4.12), so that we can observe the reason behind the changes.

## 4.4 Differential cell adhesion

The results reproduced here were obtained using Programs 19, 20, and 21.

In gastruloids, as cells start differentiating, cell states seem to sort and form an outer layer that surrounds the aggregate. This is thought to be a consequence of differential affinity between cells depending on their state. We visualize this process and show a succinct proportion dynamics analysis.

Working with low protrusion strength, we tune the adhesion factor so that cells

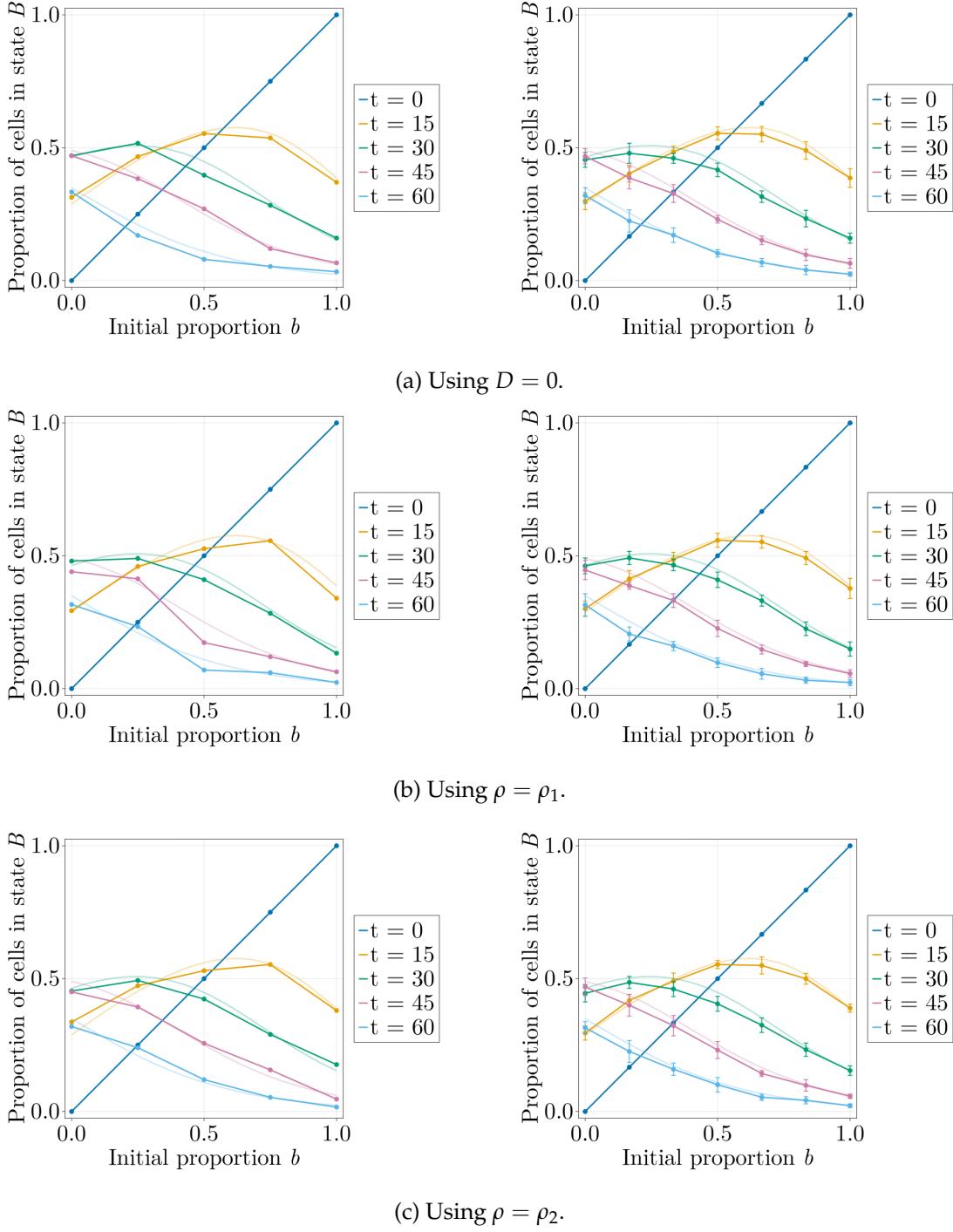


Figure 4.11:  $\phi_B(t_n)$  against  $b$  using cell-cell feedback varying  $\rho$ , and comparing a single realization (left) to the average of 10 realizations (right). Dashed lines correspond to the mean field solution, and bars indicate the standard deviation.

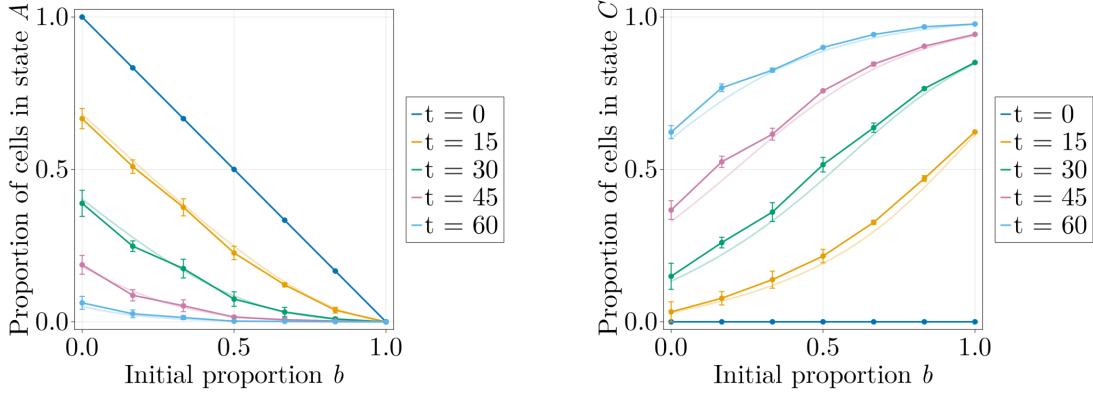


Figure 4.12:  $\phi_A(t_n)$  and  $\phi_C(t_n)$  against  $b$  using cell-cell feedback and  $\rho = \rho_1$ . Averaged over 10 realizations. Dashed lines correspond to the mean field solution, and bars indicate the standard deviation.

in state  $B$  are more attracted to each other than to the rest by a factor of  $k_B$ ,

$$\alpha_{\text{adh}}(i, j) = k_B(\mathbb{1}_B(i)\mathbb{1}_B(j)) = \begin{cases} 5 & \text{if } \text{state}(i) = \text{state}(j) = k_B \\ 0 & \text{otherwise.} \end{cases}$$

We perform a parameter sweep for  $k_B$  and determine to set it to 3-5 for the effect to be noticeable. A realization showing the outer layer of  $C$  cells is shown in Figure 4.13. We also plot the plot of  $\phi_B(t_n)$  against  $b$ , but fail to observe any significant differences in the proportions to the previous results (see Figure 4.14).

#### 4.4.1 Differential proliferation

To conclude, we consider the case in which cells of a certain state have a faster average division time than the rest. The proportions over time for these cases are presented in Figure 4.15, computed using the same parameters for proliferating while differentiating as in Section 4.2.2. The case with no proliferation (see Figure 4.15a) yields a similar result to the cell-cell signalling plots seen until now. However, the cases with proliferation evidence some differences.

In Figures 4.15b and 4.15d, the proportion of  $B$  cells starts larger than the mean field solution and then rapidly decreases. Figure 4.15c is characterized by the high proportion of  $A$  cells at the end of the differentiation.

The differentiation process of a proliferating aggregate when  $A$  cells divide faster is presented in Figure 4.16.

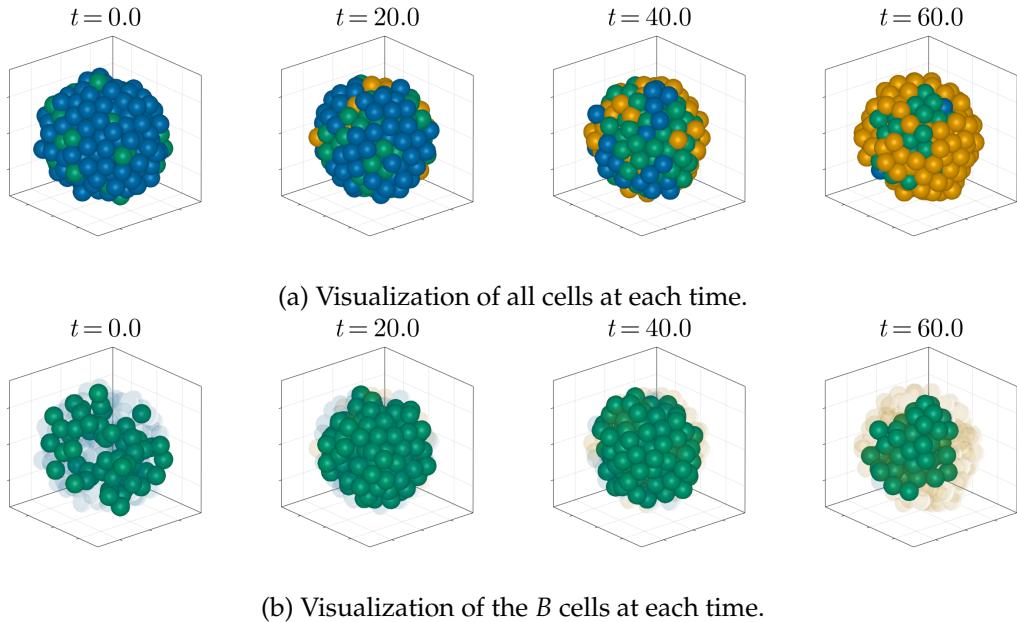


Figure 4.13: Simulation of the differentiation process using differential adhesion and cell-cell feedback, for  $k_B = 4$ . Time in hours.

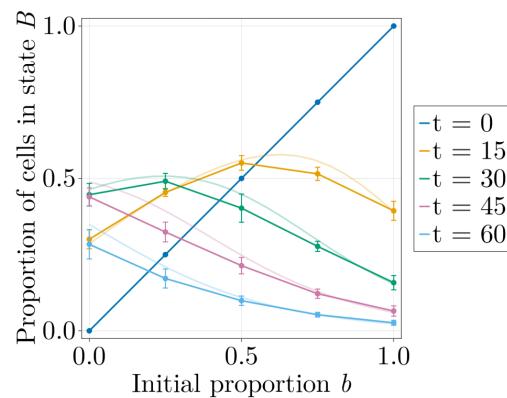


Figure 4.14:  $\phi_B(t_n)$  against  $b$  using differential adhesion,  $k_B = 5, \rho = \rho_1$ . Averaged over 8 realizations. Dashed lines correspond to the mean field solution, and bars indicate the standard deviation.

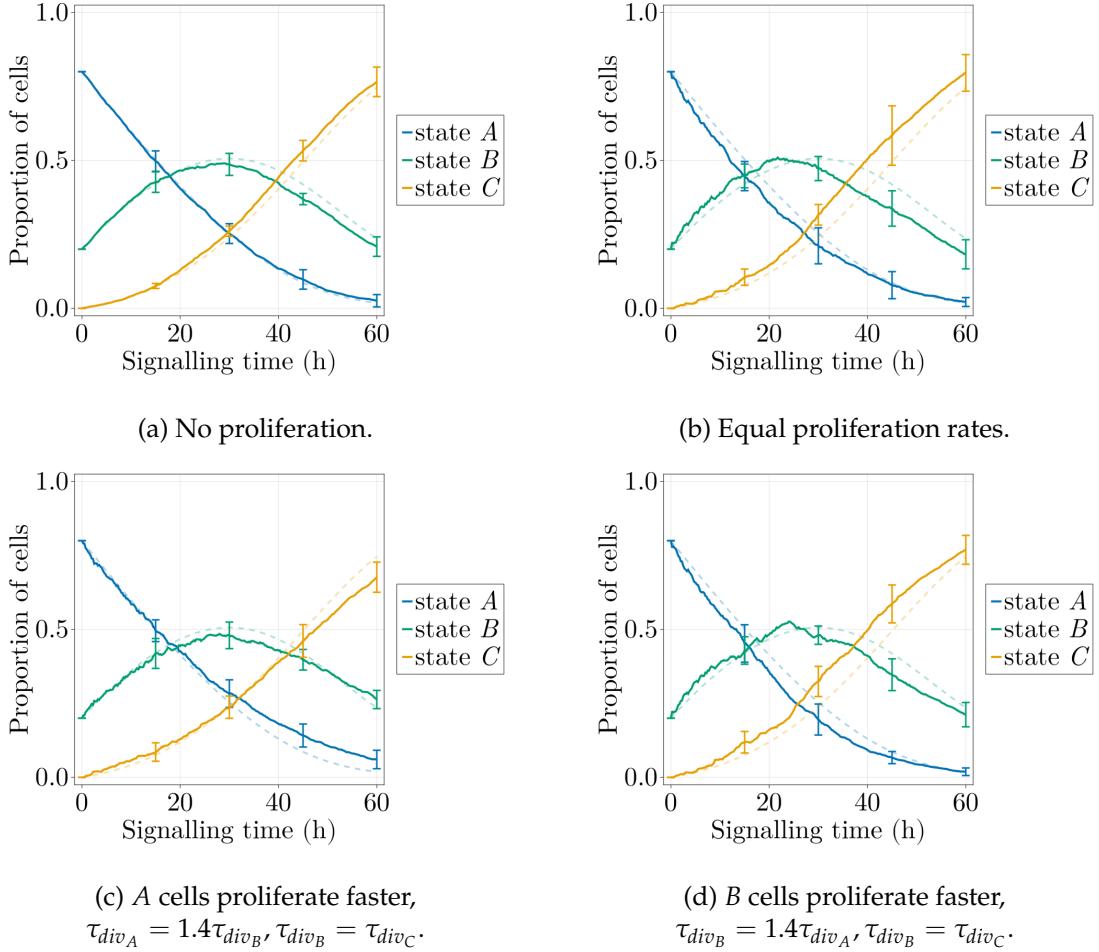


Figure 4.15: Proportions over time using differential adhesion and different proliferation rates, for  $\tau_{div} = 20, k_B = 5, \rho = \rho_1$ . Averaged over 7 realizations. Dashed lines correspond to the mean field solution, and bars indicate the standard deviation.

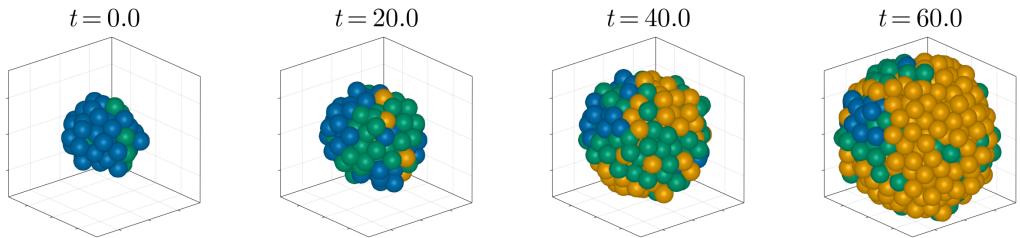


Figure 4.16: Differentiation proliferating, using differential adhesion and assuming  $A$  cells proliferate faster, for  $\tau_{div} = 20, k_B = 5, \rho = \rho_1$ .

# Chapter 5

## Conclusions and further work

In the present thesis, we have described and successfully implemented a mathematical model that simulates the early stages of embryonic organoid development, and used it to simulate observed phenomena. The model describes the movement of each one of the cells in a multicellular aggregate, and the differentiation process that these cells simultaneously go through, coupling both processes. We have discussed every step of the construction of the model in an attempt to convey the reasoning behind each decision.

The *in vitro* procedure studied consists in seeding stem cells into a well, letting them aggregate, and studying the evolution of the cell culture. Under the appropriate conditions, the aggregate eventually resembles an organ-like structure (organoid henceforth). In the model, the initial aggregate is obtained by considering the proliferation of stem cells that divide for a given time, which arises structures as those resulting from the aggregation in the well.

The study focuses on the differentiation kinetics of the aggregate – the transition of cells from one state to another. Transition can be affected by the surroundings of a cell, so we need an accurate description of the position of the cells.

We have built the equations of motion for each cell using a centre-based approach where each cell is represented as a sphere, following the models suggested by Van Liedekerke et al. 2015 and Saiz et al. 2020, accounting for passive forces, active forces and friction forces. We have found that a direct implementation of the initial proposed equations of motion leads to numerical instabilities and unphysiological aggregate configurations. Consequently, we have presented simplifications to the equations, analysed the issues encountered, and determined a final functional model for the motion in which the repulsive and

attractive passive forces follow different expressions. We have also described a method for the implementation of active forces modelled as filopodia that generate random cell-cell interactions, based on Oriola, Marin-Riera, et al. 2022.

The differentiation process has been modelled as three-state unidirectional stochastic process, where cells transition from the initial state *A* to *B*, and from state *B* to the final state *C*. Experimental data from gastruloids built using mouse embryonic stem cells (mESCs) shows that the cells present in the early stages of the formation process can be classified into three states in terms of their gene expression: T+ (*A*), T- (*B*), and further states. Results show that cells in the initial state inhibit the transition of their neighbours, and the role of the initial T- population in early development (Oriola, Torregrosa-Cortés, et al. *unpublished*).

We have implemented cell differentiation considering three possible differentiation kinetics: linear, mean field feedback, and cell-cell signalling feedback. Even though the first two can be solved analytically, cell-cell signalling is a process that depends on the ever-changing neighbours of a cell. This simulation offers a new tool for the accurate reproduction of this process, and compares it to the aforementioned simpler models.

The programming of the model has been done using CellBasedModels.jl, the Julia package for agent-based modelling, along with the provided documentation (see Torregrosa-Cortés et al. *unpublished*). Using this software, we have been able to simulate in the span of minutes processes that take days to develop in the laboratory. Also, we have contributed to expanding the documentation for this novel framework, providing examples and functions to be extrapolated to future programs.

Finally, we have tested the model using experimental measurements to simulate the cell-cell signalling feedback described in Oriola, Torregrosa-Cortés, et al. *unpublished*, and have obtained data for the evolution of the proportions over time and depending on the initial proportion of T+ cells. These simulations include varying proliferation rates and simulating the differential adhesion phenomena that lead to the formation of an outer layer in the aggregate. We conclude that the model successfully simulates the early stages of embryonic organoids, and is able to couple differentiation behaviours to the motion of cells in the aggregate.

There are several ways to expand the work presented in this text. Following this study, alternative cell-cell signalling models could be implemented. The model could be easily modified to account for factors that might be relevant

for other processes, such as cellular death, more states, or non-unidirectional transitions. Additionally, greater computation power or alternative computing methods could ease the study of computationally expensive simulations, such as proliferation during differentiation and large protrusion strength, and allow computing more realizations. Finally, a deeper statistical analysis could be performed.

# Appendix

## A.1 Mathematical derivations

**Proposition** (Proposition 2). *When the transition rates are constant, we can compute the analytical solution of Equation 2.16, depending on the values of  $p$  and  $q$ , is as follows.*

○ If  $p \neq q$ , the solution is

$$\begin{cases} \phi_A = ae^{-pt} \\ \phi_B = \frac{ap}{q-p} (e^{-pt} - e^{-qt}) + (1-a)e^{-qt} \\ \phi_C = \frac{a}{q-p} (-qe^{-pt} + pe^{-qt}) - (1-a)e^{-qt} + 1 \end{cases}$$

○ If  $p = q$ , the solution is

$$\begin{cases} \phi_A = ae^{-pt} \\ \phi_B = e^{-pt}(apt + 1 - a) \\ \phi_C = e^{-pt}(-1 - apt) + 1 \end{cases}$$

*Proof.*

1. The expression for  $\phi_A$  is solved as follows,

$$\begin{aligned} \int \frac{d\phi_A}{\phi_A} &= \int -p dt, \\ \ln(\phi_A) &= -pt + C, \\ \phi_A &= Ke^{-pt}. \end{aligned}$$

Imposing  $\phi_A(0) = a$ , we get  $K = a$ .

2. To solve  $\phi_B$ , we first multiply the differential equation by the integrating factor  $e^{qt}$ ,

$$\begin{aligned} e^{qt}\dot{\phi}_B + e^{qt}q\phi_B &= pe^{qt}\phi_A, \\ \frac{d}{dt}(e^{qt}\phi_B) &= pe^{qt}\phi_A, \\ e^{qt}\phi_B &= \int pe^{qt}\phi_A dt, \\ \phi_B &= pe^{-qt} \int ae^{(q-p)t} dt. \end{aligned}$$

○ If  $p \neq q$ ,

$$\begin{aligned} \phi_B &= ape^{-qt} \int e^{(q-p)t} dt \\ &= ape^{-qt} \left( \frac{1}{q-p} e^{(q-p)t} + C \right) \\ &\stackrel{\phi_B(0)=1-a}{=} ape^{-qt} \left( \frac{1}{q-p} e^{(q-p)t} + \frac{1-a}{ap} - \frac{1}{q-p} \right) \\ &= \frac{ap}{q-p} (e^{-pt} - e^{-qt}) + (1-a)e^{-qt}. \end{aligned}$$

○ If  $p = q$ ,

$$\begin{aligned} \phi_B &= ape^{-pt} \int e^{(p-p)t} dt = ape^{-pt}(t+C) \\ &\stackrel{\phi_B(0)=1-a}{=} e^{-pt}(apt+1-a). \end{aligned}$$

3. Solving for  $\phi_C$  reads

$$\phi_C = q \int \phi_B dt.$$

○ If  $p \neq q$ , then

$$\begin{aligned} \phi_C &= \frac{qap}{q-p} \int (e^{-pt} - e^{-qt}) dt + q(1-a) \int e^{-qt} dt \\ &= \frac{qap}{q-p} \left( -\frac{e^{-pt}}{p} + \frac{e^{-qt}}{q} \right) - (1-a)e^{-qt} + C \\ &\stackrel{\phi_C(0)=0}{=} \frac{a}{q-p} (-qe^{-pt} + pe^{-qt}) - (1-a)e^{-qt} + 1. \end{aligned}$$

○ If  $p = q$ , then

$$\begin{aligned}\phi_C &= p \int e^{-pt} (apt + 1 - a) dt \\ &= pa \int e^{-pt} ptdt + p(1-a) \int e^{-pt} dt \\ &= pa \left( -e^{-pt} t - \frac{e^{-pt}}{p} \right) - (1-a)e^{-pt} + C \\ &\stackrel{\phi_C(0)=0}{=} e^{-pt} (-1 - apt) + 1.\end{aligned}$$

□

**Proposition 4** (Proposition 3). *The dimensionless equations of motion are*

$$\begin{aligned}\tilde{\lambda} \tilde{v}_i &= \sum_{j=1}^N \tilde{F}_{ij} + P_k \tilde{F}_{p_i} \\ \frac{d\tilde{x}_i}{d\tilde{t}} &= \tilde{v}_i,\end{aligned}\tag{2.23}$$

with initial conditions

$$\tilde{v}_i(0) = \tilde{v}_{0_i}, \quad \tilde{x}_i(0) = \tilde{x}_{0_i}.\tag{2.23}$$

*Proof.*

1. Substituting the previous expressions into the first equation of 2.12 leads to

$$\left( \tilde{\lambda} \frac{F_0 T_0}{R_0} \right) \left( \tilde{v}_k \frac{R_0}{T_0} \right) = \sum_{j=1}^N F_0 \tilde{F}_{ij} + F_0 P_i \tilde{F}_p,$$

and constants cancel out leading to the described equation.

2. Substituting the previous expressions into the second equation of 2.12 leads to

$$\frac{d(\tilde{x}_i R_0)}{d(\tilde{t} T_0)} = \tilde{v}_i \frac{R_0}{T_0},$$

and constants cancel out leading to the described equation.

□

## A.2 Code

### A.2.1 Technical specifications

Computations have been carried out using the system configuration described in Table 5.1.

Device	<i>Lenovo Ideapad 5-15ITL05 (82FG)</i>
OS	<i>Linux Mint 22 Cinnamon (6.2.9)</i>
Julia	<i>Julia 1.9.4+0.x64.linux.gnu</i>
CBM.jl	<i>CellBasedModels v0.1.0</i>

Table 5.1: Technical specifications.

### A.2.2 Programs

All the code has been written in Julia using Jupyter Notebooks unless stated otherwise.

**Program 1** (preamble/packages). *Loads the packages needed for the project.*

**Program 2** (preamble/model). *Loads the global friction model (see Equation 2.23).*

**Program 3** (preamble/functions). *Loads most of the functions used by the programs.*

**Program 4** (3.01). *Visualization of the issue regarding the increase of neighbours. Averaged over several realizations.*

**Program 5** (3.02). *Analysis of the issue regarding the increase of neighbours.*

**Program 6** (3.03). *Parameter sweep for  $\alpha_{adh}$  using the variable friction model.*

**Program 7** (3.04). *Parameter sweep for  $\alpha_{adh}$  using the constant friction model*

**Program 8** (3.05). *Proposed model with constant fiction.*

**Program 9** (3.06). *Instability of the model using Equation 3.10.*

**Program 10** (3.07). *Computation of the sum of velocities of the neighbours using the model introduced in Program 6.*

**Program 11** (4.00). *Example displaying the main features and functions of the code.*

**Program 12** (4.01). *Loop that simulates communities using different parameters in order to perform parameter sweep and ensure stability.*

**Program 13** (4.02). *Simulation assuming that proliferation continues during differentiation.*

**Program 14** (4.03). *Measurement of the growth of cells.*

**Program 15** (4.04). *Measurement of the effect of protrusions.*

**Program 16** (4.05). *Average of the proportions of cell types for several realizations and comparison of the output to known solutions. It includes computations for several configurations.*

**Program 17** (4.06). *Replica of the computations performed in Program 16 using the set-up described in Program 13.*

**Program 18** (4.07 and 4.08). *Computation of the proportion of cells in state B in terms of its initial proportion for a given set of timestamps, and comparison with known solutions. For a single realization in 4.07, and averaged in 4.08.*

**Program 19** (4.09). *Simulation assuming that force between two cells can change depending on their states. Plots the cells in each state to visualize the effect.*

**Program 20** (4.10). *Replica of the computations performed in Programs 16 and 18 using the set-up described in Program 19.*

**Program 21** (4.11). *Replica of the computations performed in Programs 16 using the set-up described in Program 19 along with proliferation during the differentiation process.*

**Program 22** (`images2gif.py`). *Python script that finds plots in the folder and converts them into a GIF.*

### A.2.3 Packages

The main package, `CellBasedModels.jl`, already includes some features such as numerical integrators and probability distributions. However, some tasks require extra packages.

- All plots are created using `GLMakie` (Danisch and Krumbiegel 2021), and formatted with `MathTeXEngine`. The colours used are from the colour-blind friendly palette presented in Wong 2011.
- Sampling and random seeds are provided by `Distributions` and `Random`, respectively.

- File formatting is handled by the packages `Printf` and `Dates`, and data extraction by `Glob`.
- Preamble files are included in the program using `NBInclude`.
- The Python script (see Program 22) requires the `imageio` package.

### A.3 Additional figures

Figure 5.1 corresponds to the proportion of cells in states  $A$  and  $C$  in terms of  $\phi_B(0) = b$ . The analogous plot for  $B$  is studied in section 4.3.

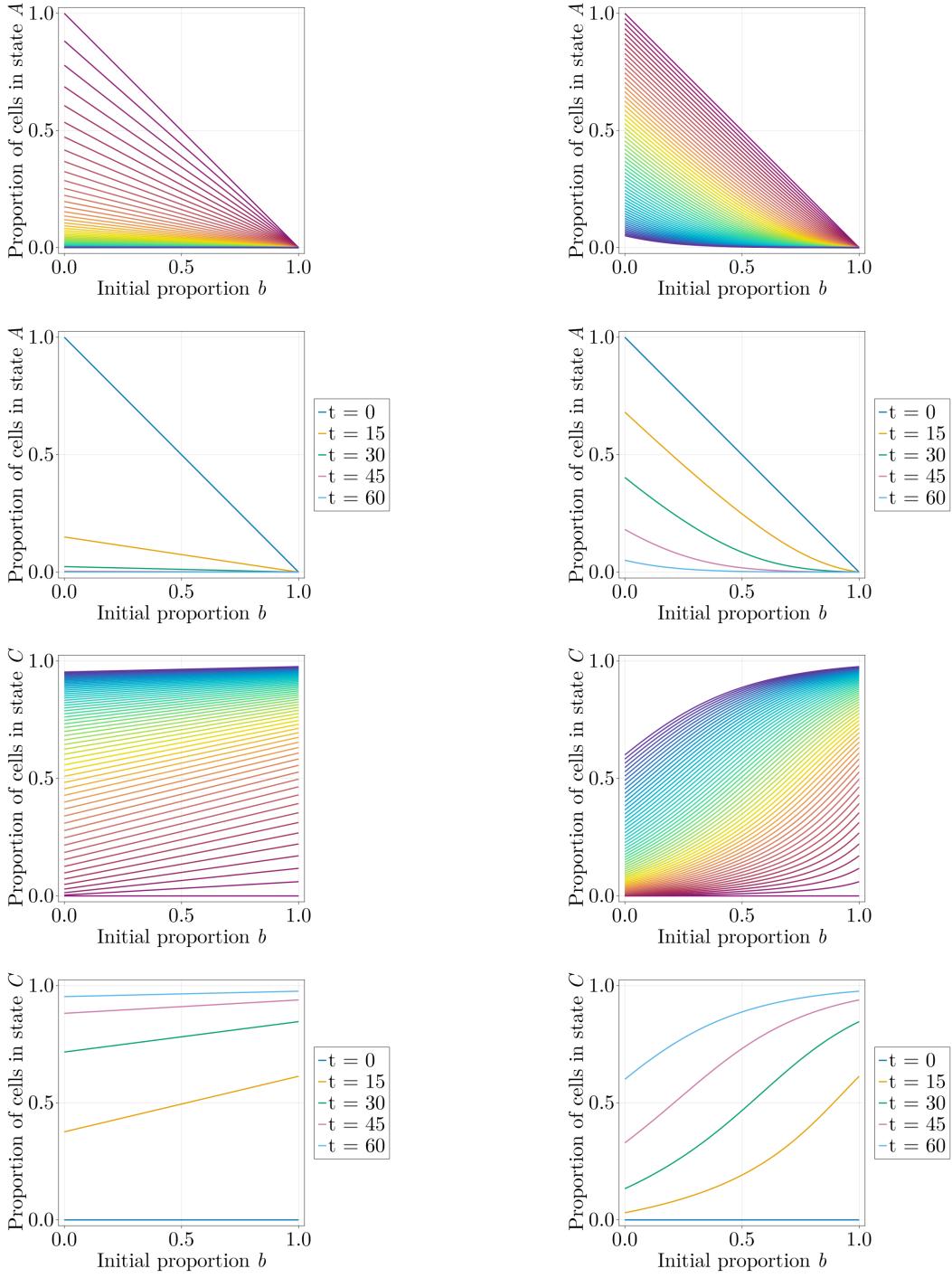


Figure 5.1:  $\phi_X(t_n)$  against  $b$  using the known solutions,  $X = A, B$  for constant rates (left) and mean field feedback (right). Plots for each second  $s \in [0, 60]$  when unspecified.

# References

- Alberts, Bruce, Alexander Johnson, and et al. Lewis Julian (2002). *Molecular Biology of the Cell. 4th edition.* London, England: Garland Science. URL: <https://www.ncbi.nlm.nih.gov/books/NBK26939/>.
- Bezanson, Jeff et al. (2017). "Julia: A Fresh Approach to Numerical Computing". In: *SIAM Review* 59.1, pp. 65–98. DOI: [10.1137/141000671](https://doi.org/10.1137/141000671).
- Brink, Susanne C. van den et al. (2014). "Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells". In: *Development* 141.22, pp. 4231–4242. ISSN: 0950-1991. DOI: [10.1242/dev.113001](https://doi.org/10.1242/dev.113001).
- Bronson, Stuart (2022). *Predictor-corrector methods and Runge-Kutta*. Accessed 2024-09-22. URL: <https://math.libretexts.org/>.
- Cooke, Christopher B. et al. (2023). "Gastruloid-derived primordial germ cell-like cells develop dynamically within integrated tissues". In: *Development* 150.17, dev201790. ISSN: 0950-1991. DOI: [10.1242/dev.201790](https://doi.org/10.1242/dev.201790).
- Danisch, Simon and Julius Krumbiegel (2021). "Makie.jl: Flexible high-performance data visualization for Julia". In: *Journal of Open Source Software* 6.65, p. 3349. DOI: [10.21105/joss.03349](https://doi.org/10.21105/joss.03349).
- Germann, Philipp, Miquel Marin-Riera, and James Sharpe (2019). "ya || a: GPU-powered Spheroid Models for Mesenchyme and Epithelium". In: *bioRxiv*. DOI: [10.1101/525352](https://doi.org/10.1101/525352). URL: <https://www.biorxiv.org/content/early/2019/01/20/525352>.
- Gritti, Nicola, David Oriola, and Vikas Trivedi (2021). "Rethinking embryology in vitro: A synergy between engineering, data science and theory". In: *Developmental Biology* 474. Synthetic Embryology, pp. 48–61. ISSN: 0012-1606. DOI: [10.1016/j.ydbio.2020.10.013](https://doi.org/10.1016/j.ydbio.2020.10.013).
- Huch, Meritxell et al. (2017). "The hope and the hype of organoid research". In: *Development* 144.6, pp. 938–941. ISSN: 0950-1991. DOI: [10.1242/dev.150201](https://doi.org/10.1242/dev.150201).
- Liu, Allen P and Daniel A Fletcher (2009). "Biology under construction: in vitro reconstitution of cellular function". In: *Nature Reviews Molecular Cell Biology* 10.9, pp. 644–650. DOI: [10.1038/nrm2746](https://doi.org/10.1038/nrm2746).

- Moore, Simon W., Pere Roca-Cusachs, and Michael P. Sheetz (2010). "Stretchy Proteins on Stretchy Substrates: The Important Elements of Integrin-Mediated Rigidity Sensing". In: *Developmental Cell* 19.2, pp. 194–206. ISSN: 1534-5807. DOI: [10.1016/j.devcel.2010.07.018](https://doi.org/10.1016/j.devcel.2010.07.018).
- Oriola, David, Miquel Marin-Riera, et al. (2022). "Arrested coalescence of multicellular aggregates". In: *Soft Matter* 18, p. 3771. DOI: [10.1039/d2sm00063f](https://doi.org/10.1039/d2sm00063f).
- Oriola, David, Gabriel Torregrosa-Cortés, et al. (unpublished). "Cell-cell communication controls the timing of symmetry-breaking in an embryo-like system".
- Pillarisetti, Anand et al. (2009). "Mechanical characterization of mouse embryonic stem cells". In: *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1176–1179. DOI: [10.1109/IEMBS.2009.5333954](https://doi.org/10.1109/IEMBS.2009.5333954).
- Purcell, E. M. (1977). "Life at low Reynolds number". In: *American Journal of Physics* 45.1, pp. 3–11. ISSN: 0002-9505. DOI: [10.1119/1.10903](https://doi.org/10.1119/1.10903).
- Roccio, Marta et al. (2013). "Predicting stem cell fate changes by differential cell cycle progression patterns". In: *Development* 140.2, pp. 459–470. ISSN: 0950-1991. DOI: [10.1242/dev.086215](https://doi.org/10.1242/dev.086215).
- Rosen, Mary Ellen, Christopher P Grant, and J C Dallon (2021). "Mean square displacement for a discrete centroid model of cell motion". en. In: *PLoS One* 16.12, e0261021. DOI: [10.1371/journal.pone.0261021](https://doi.org/10.1371/journal.pone.0261021).
- Saiz, Néstor et al. (2020). "Growth-factor-mediated coupling between lineage size and cell fate choice underlies robustness of mammalian development". In: *eLife* 9, e56079. DOI: [10.7554/eLife.56079](https://doi.org/10.7554/eLife.56079).
- Serra, Denise et al. (2019). "Self-organization and symmetry breaking in intestinal organoid development". In: *Nature* 569.7754, pp. 66–72. DOI: [10.1038/s41586-019-1146-y](https://doi.org/10.1038/s41586-019-1146-y).
- Simunovic, Mijo and Ali H. Brivanlou (2017). "Embryoids, organoids and gastruloids: new approaches to understanding embryogenesis". In: *Development* 144.6, pp. 976–985. ISSN: 0950-1991. DOI: [10.1242/dev.143529](https://doi.org/10.1242/dev.143529).
- Süli, Endre and David F. Mayers (2003). *An Introduction to Numerical Analysis*. Cambridge University Press. ISBN: 9780521810265. DOI: [10.1017/cbo9780511801181](https://doi.org/10.1017/cbo9780511801181).
- Sullivan, Adrienne E. and Silvia D. M. Santos (2023). "The ever-growing world of gastruloids: autogenous models of mammalian embryogenesis". In: *Current Opinion in Genetics & Development* 82, p. 102102. ISSN: 0959-437X. DOI: [10.1016/j.gde.2023.102102](https://doi.org/10.1016/j.gde.2023.102102).
- Torregrosa-Cortés, Gabriel (2023). "Dynamics of cellular decision making in embryogenesis". PhD thesis. Barcelona: Universitat Pompeu Fabra. URL: <http://hdl.handle.net/10803/689098>.

- Torregrosa-Cortés, Gabriel et al. (unpublished). "CellBasedModels.jl: A flexible and efficient cell-based modeling framework in Julia". URL: <https://github.com/dsb-lab/CellBasedModels.jl>.
- Turner, David A. (2017). "From Organoids to Gastruloids". In: *The Biologist* 64.5, pp. 14–19. URL: <https://www.rsb.org.uk/biologist-features/from-organoids-to-gastruloids>.
- Van Liedekerke, Paul et al. (2015). "Simulating tissue mechanics with agent-based models : concepts, perspectives and some novel results". In: *Computational Particle Mechanics* 2.4, pp. 401–444. ISSN: 2196-4378. DOI: [10.1007/s40571-015-0082-3](https://doi.org/10.1007/s40571-015-0082-3).
- Wong, Bang (2011). "Points of view: Color blindness". In: *Nature Methods* 8.6, pp. 441–441. DOI: [10.1038/nmeth.1618](https://doi.org/10.1038/nmeth.1618).
- Yusko, Erik C. and Charles L. Asbury (2014). "Force is a signal that cells cannot ignore". In: *Molecular Biology of the Cell* 25.23. PMID: 25394814, pp. 3717–3725. DOI: [10.1091/mbc.e13-12-0707](https://doi.org/10.1091/mbc.e13-12-0707).